View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by INRIA a CCSD electronic archive server

HAL archives-ouvertes.fr

Computable Measure Theory and Algorithmic Randomness

Mathieu Hoyrup, Jason Rute

To cite this version:

Mathieu Hoyrup, Jason Rute. Computable Measure Theory and Algorithmic Randomness. Handbook of Computable Analysis, In press. hal-02938919

HAL Id: hal-02938919

https://hal.inria.fr/hal-02938919

Submitted on 15 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Computable Measure Theory and Algorithmic Randomness

Mathieu Hoyrup and Jason Rute

Abstract We provide a survey of recent results in computable measure and probability theory, from both the perspectives of computable analysis and algorithmic randomness, and discuss the relations between them.

1 Introduction

The underlying topic of this chapter is computable probability theory, considered from two angles. The first angle follows the traditional approach of computable analysis, where the goal is to develop effective versions of classical notions and to study the effectiveness of classical theorems. These notions and theorems mainly come from measure theory in which probability theory is grounded. Therefore, the first part of this chapter is devoted to computable measure theory. The second direction is the algorithmic theory of randomness, whose goal was originally to define what it means for an individual object to be random, using computability theory. These two approaches have been developed in parallel for a long time, but their interaction has recently become a fruitful research direction, surveyed in this chapter.

Let us give a quick overview of the typical questions that are studied in the vast area of computable probability theory. A recurrent topic in computable analysis is to investigate the effectiveness of existence theorems. Many theorems in probability theory are convergence theorems. How to analyze such theorems from a computability perspective? A theorem stating the convergence of a sequence can be presented as an existence theorem: it asserts the existence of an index from which the terms of

Mathieu Hoyrup
Université de Lorraine, CNRS, Inria, LORIA, F 54000 Nancy, France, e-mail: mathieu.hoyrup@inria.fr

Jason Rute
Department of Mathematics Pennsylvania State University State College, PA 16802, e-mail: jason.rute@uwalumni.com

the sequence are close to the limit. This leads to investigating the computability of the speed of convergence, i.e. of the aforementioned index given the prescribed distance to the limit. It happens that many convergence theorems are not computable in this way (for instance, martingales convergence theorems or ergodic theorems). This was observed by Bishop in the context of constructive analysis, in his Foundations of Constructive Analysis (p. 214 in [19]):

> Certain parts of measure theory are hard to develop constructively, because limits that are classically proved to exist simply do not exist constructively.

So it looks like it is the end of the story. However there is another way of interpreting an almost-sure convergence theorem as an existence theorem: it states the existence of a set of full measure on which pointwise convergence holds. So the computability problem amounts to studying the computability of this full measure set.

Some of the computable approaches to measure theory fail to provide useful computability notions for full measure sets: the full measure sets are all computable because they are all equivalent to the whole space. One needs a finer look into the computability of those sets, and Algorithmic Randomness provides this.

Algorithmic Randomness starts with Martin-Löf's seminal paper [85] where he introduces a notion of effective null set, nowadays called *Martin-Löf null set*. Such a notion enables one to define what a random point is: it is a point that does not belong to any effective null set, which makes sense as there are countably many such sets hence their union is again a null set. Since then Algorithmic Randomness has been studied in several directions. We will only present the part that interacts with computable measure and probability theory and will not mention its interactions with computability theory and Kolmogorov complexity, a large part of which can be found in [82, 95, 32].

The goals of this chapter are to present the main results obtained in recent years, a comprehensive bibliography as well as the basic definitions, tools and techniques needed for this development. We include proofs of simple results, some of them appearing nowhere explicitly, and so that the reader can quickly understand how they work and confidently use them. As for the deeper and more complicated results, we refer to the corresponding research articles, where the proofs can be found. When possible, we give an outline of the proof, or at least some intuition about how it works.

2 Computable measure theory

2.1 Background from computable analysis

We assume familiarity with basic notions from computability theory (computable set of natural numbers, computably enumerable (c.e.) set, computable function, etc.). We give a couple of central notions from computable analysis. References will be given where needed. The standard reference for computability on countably-based spaces is [116].

A ***computable metric space*** is a triple (X,d,S) where (X,d) is a separable metric space and $S=(s_i)_{i\in\mathbb{N}}$ is a sequence of points of X, called *simple* points, such that the reals numbers $d(s_i,s_j)$ are uniformly computable. A ***name*** of a point $x\in X$ is a function $f:\mathbb{N}\to\mathbb{N}$ such that $d(x,s_{f(i)})<2^{-i}$ for all i. A point is ***computable*** if it has a computable name. The basic metric balls are the metric balls centered at simple points with rational radii, and have a canonical indexing $(B_i)_{i\in\mathbb{N}}$. An open set $U\subseteq X$ is an ***effective open set*** if $U=\bigcup_{i\in E}B_i$ for some c.e. set $E\subseteq\mathbb{N}$. A function $f:X\to Y$ between computable metric spaces is computable if there is an oracle Turing machine reading a name of x and outputting a name of $f(x)$. Equivalently, f is computable if for each basic ball $B\subseteq Y$ the pre-image $f^{-1}(B)$ is an effective open subset of X, uniformly in the index of B.

A compact set $K\subseteq X$ is an ***effective compact set*** if there is a computable enumeration of the finite sets $F\subseteq\mathbb{N}$ such that $K\subseteq\bigcup_{i\in F}B_i$. Let $K\subseteq X$ be an effective compact set. Its complement $X\setminus K$ is effectively open. If U is effectively open then $K\setminus U$ is effectively compact. If $f:K\to Y$ is computable then $f(K)$ is effectively compact. If f is moreover one-to-one then $f^{-1}:f(K)\to K$ is computable.

For instance, $\mathbb{R}$ with the Euclidean metric and a canonical enumeration $(q_i)_{i\in\mathbb{N}}$ of the rational numbers is a computable metric space. We denote by $\mathbb{R}_<$ the set of real numbers with a different naming system, or representation. In that space, a name of $x\in\mathbb{R}_<$ is a function $f:\mathbb{N}\to\mathbb{N}$ such that $x=\sup_i q_{f(i)}$. A real number is ***lower semicomputable*** or ***left-c.e.*** if it has a computable name in this sense.

2.2 Framework

The most general way of defining measures goes through the abstract notions of ring, algebra, σ-ring, σ-algebra, outer measure, etc. Effective counterparts to this development have been investigated and used in several articles [120, 121, 69, 70, 119, 118].

Another approach, less general but covering a wide range of applications, is to restrict oneself to spaces with a structure (for instance a separable metric), on which computable analysis is already settled, and to work with the measurable structure induced by it (for instance the Borel σ-algebra, generated by the open sets). Many articles in the literature follow this approach, which we will adopt here for the following reasons.

The practical reason is that manipulating a computable measurable structure, in addition to other structures, is rather heavy. The mathematical reason is that results from ordinary measure and probability theory are often stated on spaces with more structure than just a measurable structure, like a metric for instance, and hold only on such spaces. As a result, the computable structures from mainstream computable analysis are usually sufficient to investigate computable measure theory. Finally, the approach of working with a fixed ring shows its limits when the measure is not fixed and computable but is a variable object of the problem. This limitation manifests itself in several ways. Each ring induces its own notion of computable measure, even among the computable rings, so the underlying ring should evolve at the same time as the measure. The topology induced by the representation of measures associated with a fixed ring is not the important weak* topology. Intuitively, fixing a ring and representing a measure by giving the weights of the ring elements introduces artificial discontinuities, similar to representing real numbers by their binary expansions.

For all these reasons we have chosen to follow the second approach by working on computable metric spaces only and considering Borel measures. It avoids too many additional definitions and it is sufficient for most purposes. Moreover we will restrict our attention to probability measures for simplicity.

We are implicitly working in the framework of Type-Two computability and represented spaces, but there other options. A complete approach to computable measure theory has been developed by Edalat [33, 34] in the framework of domain theory.

2.2.1 Representing and computing with probability measures

Let X be a computable metric space. A Borel probability measure can be equivalently represented as (see [49, 107] for instance) :

- A function $\mathscr{O}(X) \to \mathbb{R}_<$ mapping an open set U to $\mu(U)$,
- A function $\mathscr{C}(X,[0,1]) \to \mathbb{R}$ mapping a bounded continuous function f to $\int f \, \mathrm{d}\mu$,
- A function $\mathbb{N} \to \mathbb{R}_<$ mapping an index of a finite union of balls to its weight,
- A point in the computable metric space $\mathscr{M}_1(X)$ of Borel probability measures endowed with the Prokhorov metric.

Here $\mathbb{R}$ denotes the space of real numbers with the Euclidean topology while $\mathbb{R}_<$ denotes the space of real numbers with the topology induced by the semi-lines $(x,+\infty)$.

In particular,

Definition 2.1 (Computable measure). A Borel probability measure μ is ***computable*** if the following equivalent conditions hold:

- The measure of a finite union of basic metric balls is uniformly lower semicomputable,
- The measure of an effective open set is uniformly lower semicomputable,
- The integral of a bounded computable function $h : X \to [0,1]$ is uniformly computable.

More generally, a finite measure μ is computable if $\mu(X)$ is a computable real number and the normalized measure $\mu/\mu(X)$ is a computable probability measure. Computability of σ-finite or general measures can be defined in similar ways, with variations.

If μ is a computable Borel probability measure then the measure of a basic ball $B(s,r)$ is lower semicomputable but is not necessarily computable. However one can take a class of radii other than the rationals to make the measures of basic balls computable. This is done by ensuring that $\mu(\overline{B}(s,r) \setminus B(s,r)) = 0$ [21]. These new balls are called μ-continuity balls.

A computable Riesz representation theorem is proved in [84, 71]: a measure can be equivalently represented as the positive linear functional mapping continuous real-valued functions with compact support to their integrals. The result is proved for locally compact Hausdorff spaces satisfying computability assumptions, and Borel regular measures.

A measure can be alternatively represented as a valuation on the lower semicontinuous functions [33, 107].

2.2.2 Effectively approximable sets and functions

The notion of a computable function expresses the intuitive idea of an algorithm processing an input and producing an output. Computable functions being necessarily continuous, this notion is sometimes too restrictive and one needs a weaker notion, based on the idea that the algorithm performing the computation is allowed to make mistakes on a small set of inputs, in a controlled way. This is the motivation underlying the notion of an effectively approximable function. This definition was introduced by Ko [79] on Euclidean spaces and generalized to other topological spaces by Bosserhoff [21].

Definition 2.2. Let X,Y be computable metric spaces and μ a computable Borel probability measure over X. A function $f : X \to Y$ is ***effectively μ-approximable*** if there exists an oracle Turing machine that given a name of $x \in X$ and a rational $\varepsilon > 0$ outputs a name of $f(x)$ for all x in a set of measure at least $1-\varepsilon$.

One may require a weaker condition: given ε, δ, the machine computes $f(x)$ at precision δ for all x in a set of measure at least $1-\varepsilon$. This apparently weaker requirement is actually equivalent to the one given in the definition and is sometimes simpler to prove. Moreover it is sufficient to check this condition only when $\delta = \varepsilon$.

Example 2.3 (Random harmonic series). If a binary sequence $s \in \{0,1\}^{\mathbb{N}}$ is obtained by independently tossing a fair coin then the sum

$$f(s) = \sum_{n\geq 1} \frac{(-1)^{s_n}}{n}$$

converges almost surely. Can the sum be computed from s? The function $s \mapsto f(s)$ is not computable because it is obviously not continuous: knowing the n first values of s gives no information about the limit, which can be any real number.

However one can easily prove that the function f is effectively approximable w.r.t. the uniform measure over $\{0,1\}^{\mathbb{N}}$. Let δ, ε be positive rational numbers. For each m, consider the random variable $T_m = \sum_{n>m}(-1)^{s_n}/n$, whose expected value is $\mathbb{E}[T_m] = 0$. Its variance $\mathbb{E}[T_m^2] = \sum_{n>m} 1/n^2$ converges effectively to 0 as m grows, so by Chebyshev's inequality $\mathbb{P}[|T_m| \geq \delta] \leq \mathbb{E}[T_m^2]/\delta^2$ can be taken as small as we want, in particular smaller than ε, by taking m sufficiently large. The sum $\sum_{n=1}^{m}(-1)^{s_n}/n$ is then δ-close to $f(s)$ for all s in a set of measure at least $1-\varepsilon$, and that finite sum can be uniformly computed from s and ε, δ.

Example 2.4 (Pólya urn). In an urn starting with one black ball and one white ball, at each round one draws a ball at random and put it back into the urn together with a new ball of the same color. The sequence of observed colors follows a probability distribution μ over $\{0,1\}^{\mathbb{N}}$ (interpreting 1 as black and 0 as white for instance) defined by

$$\mu([u]) = \frac{(|u|_0)! \cdot (|u|_1)!}{(1+|u|)!},$$

where $u \in \{0,1\}^*$, $|u|_0$ is the number of occurrences of 0 in u and $|u|_1$ is the number of occurrences of 1 in u (0! evaluates to 1 here). For μ-almost every outcome sequence $s \in \{0,1\}^{\mathbb{N}}$, the frequency of 1's in the sequence converges to a real number $p(s)$. Can $p(s)$ be computed from s? Again the function p is heavily discontinuous hence not computable, but one can prove that it is effectively μ-approximable. This can be done as in the previous example by showing that the speed of convergence is effective (this will be formalized in Section 2.2.5), or by a more abstract argument that will be presented in Examples 2.27 and 4.14.

The notion of effective approximability for functions has an immediate counterpart for sets [79].

Definition 2.5. Let X be a computable metric space and μ a computable Borel probability measure over X. A set $A \subseteq X$ is ***effectively μ-approximable*** if its characteristic function $1_A : X \to \{0,1\}$ is effectively μ-approximable.

If $A \subseteq X$ is effectively μ-approximable then $\mu(A)$ is computable.

Example 2.6 (Smith-Volterra-Cantor set). The Smith-Volterra-Cantor set, which is also known as the fat Cantor set, is a nowhere dense closed subset of $[0,1]$ that has positive Lebesgue measure. It is obtained by starting from $[0,1]$ and by removing, at each stage $n \geq 1$, subintervals of width $1/4^n$ from the middle of each remaining interval. The limit set has Lebesgue measure $1/2$. This set is effectively λ-approximable, by the next observation.

Proposition 2.7 (Ko [79]). *Let A be an effective closed set or an effective open set. A is effectively μ-approximable if and only if $\mu(A)$ is computable.*

If $f : X \to [0,+\infty]$ is effectively μ-approximable and *bounded* then $\int f \, \mathrm{d}\mu$ is computable.

If $f : X \to [0,+\infty]$ is effectively μ-approximable but is *unbounded* then $\int f \, \mathrm{d}\mu$ is not necessarily computable, but is always lower semicomputable.

Example 2.8 (Non-computable integral). Let $(n_i)_{i\in\mathbb{N}}$ be a computable sequence enumerating a non-computable set $A \subseteq \mathbb{N}$ such as the halting set. The piecewise constant function $f : [0,1] \to [0,+\infty)$ defined by $f(x) = 2^{-n_i+i+1}$ on $(2^{-i-1}, 2^{-i}]$ and $f(0) = 0$

is effectively λ-approximable (there is an algorithm that computes f outside the null set $\{0\} \cup \{2^{-i-1} : i \in \mathbb{N}\}$) but its integral is $\sum_{n \in A} 2^{-n}$ which is not a computable real number.

Composing effectively approximable functions can be done for an appropriate choice of the involved measures. A measurable function $f : X \to Y$ pushes any measure μ over X to a measure over Y, denoted by μ_f and called the push-foward of μ under f. It is defined by $\mu_f(A) = \mu(f^{-1}(A))$ for all measurable sets A.

Proposition 2.9 (Bosserhoff [21]). *If $f : X \to Y$ is effectively μ-approximable then the push-forward measure μ_f is computable. If $g : Y \to Z$ is effectively μ_f-approximable then $g \circ f$ is effectively μ-approximable.*

Proof. To prove that ν is computable, we show that if $h : Y \to [0,1]$ is a bounded computable function then $\int h \, \mathrm{d}\nu$ is computable, uniformly in h. By definition of μ_f, one has $\int h \, \mathrm{d}\mu_f = \int h \circ f \, \mathrm{d}\mu$. The function $h \circ f$ is effectively μ-approximable (simply compose the algorithms for f and h) so $\int h \circ f \, \mathrm{d}\mu$ is computable as h is bounded. Everything is uniform in h.

If g is effectively μ_f-approximable then the algorithms approximating f and g can be easily composed to approximate $g \circ f$. □

The computational complexity of integration has been investigated by Ko [79] and Kawamura [73], for continuous functions on $[0,1]$ with the Lebesgue measure. They essentially proved that the complexity of integration corresponds to the counting class #P and is in some sense complete for this class. For other measures than the Lebesgue measure, the dependence of the complexity of integration on the measure has been investigated by Férée and Ziegler [35].

2.2.3 Effective measurability

Computable functions can be seen as the effective version of continuous functions: a function f is computable if and only if it is effectively continuous in the sense that the pre-image of effective open sets are effective open sets, uniformly; also, a function is continuous if and only if it is computable relative to some oracle.

In the same way, the weaker notion of effectively approximable function can be interpreted as the effective counterpart of a classical notion, measurability. To this end, we briefly present a notion of computability and a representation for measurable sets.

Let X be a computable metric space and μ a computable Borel probability measure over X. The set of Borel subsets of X can be endowed with a pseudometric $d_\mu(A,B) = \mu(A \Delta B)$. The quotient of this pseudometric space by the equivalence relation $A \equiv_\mu B \iff d_\mu(A,B) = 0$ is a separable metric space. We now show how to choose a countable dense subset in order to make it a computable metric space.

The space X has a topological basis of sets whose measures are computable real numbers. This basis is obtained by computing a sequence of positive real numbers $(r_i)_{i \in \mathbb{N}}$ that is dense in $(0,\infty)$, such that $\mu(\overline{B}(s,r_i) \setminus B(s,r_i)) = 0$ for all i and s

in the countable dense subset associated to X (see [21]). The ring $\mathscr{R}$ obtained by taking the closure of this basis under finite unions and complements has a canonical numbering $\mathscr{R} = \{R_i : i \in \mathbb{N}\}$. One can then take $\mathscr{R}$ as a dense sequence in the quotient space of measurable sets.

The computable metric space of (equivalence classes of) measurable sets induces a representation and a computability notion for measurable sets: A is represented by any Cauchy sequence of ring elements converging fast to A in the pseudometric d_μ. An equivalent representation consists in describing the sequence of real numbers $\mu(A \cap R_i)$.

The next definition and result appeared in [79] on $\mathbb{R}$ and in [21] on general spaces.

Definition 2.10. A set A is ***effectively μ-measurable*** if the following equivalent conditions hold:

- Given a positive rational ε one can compute, uniformly in ε, a finite union A_ε of basic μ-continuity balls such that $\mu(A \Delta A_\varepsilon) \leq \varepsilon$,
- Given i one can compute $\mu(A \cap R_i)$ uniformly in i.

Proposition 2.11. *A set is effectively μ-approximable if and only if it is effectively μ-measurable.*

For a proof, one can consult [79, 21]. Other ways of representing measurable sets have been investigated on general measurable spaces with σ-finite measures [120, 119]. One can also define the notion of an effectively μ-measurable function and prove that it is equivalence to effective μ-approximability. We do not include it here as it would require some extra definitions and will not be used in this chapter. The interested reader can consult [79, 118].

2.2.4 L^p-spaces and absolute continuity

For each computable real number p, the space $L^p(X,\mu)$ is a complete separable metric space that is naturally a computable metric space. The finite linear combinations with rational coefficients of characteristic functions of elements of $\mathscr{R}$ are dense and their canonical numbering makes the metric

$$d(f,g) = \|f-g\|_p = \left(\int_X |f-g|^p \,\mathrm{d}\mu\right)^{1/p}$$

computable. This structure induces a representation of the elements in $L^p(X,\mu)$, which are equivalence classes of functions under μ-almost everywhere coincidence, and a notion of $L^p(X,\mu)$-computable function.

Definition 2.12. A measurable function $f : X \to \mathbb{R}$ is $L^p(X,\mu)$***-computable*** if its equivalence class is a computable element of $L^p(X,\mu)$.

Several characterizations of L^p-computable functions have been obtained [99, 122]. L^p-computability is not far away from effective approximability, as the next result shows (an indirect proof appears in [66]).

Proposition 2.13. *A function $f : X \to \mathbb{R}$ is $L^p(X,\mu)$-computable if and only if it is effectively μ-approximable and $\|f\|_p$ is computable.*

Any non-negative function in $L^1(X,\mu)$ induces a finite measure ν with density f w.r.t. to μ:

$$\nu(A) = \int_A f \,\mathrm{d}\mu.$$

(here we assume that $\|f\|_1 = 1$ so that ν is a probability measure). The measure ν is absolutely continuous w.r.t. μ, written $\nu \ll \mu$, which means that for every measurable set A, $\mu(A) = 0$ implies $\nu(A) = 0$. The Radon-Nikodym theorem asserts that every absolutely continuous measure can be obtained this way. The function f is called the Radon-Nikodym derivative of ν w.r.t. μ and is denoted by $\frac{d\nu}{d\mu}$.

The computability of the correspondence between absolutely continuous measures and densities in L^1 has been investigated. We present a few results in this direction.

Definition 2.14 (Effective absolute continuity). Let μ,ν be Borel probability measures. We say that ν is ***effectively absolutely continuous w.r.t.*** μ if there exists a computable function $\varphi : \mathbb{N} \to \mathbb{N}$ such that $\mu(A) \leq 2^{-\varphi(n)}$ implies $\nu(A) \leq 2^{-n}$ for all Borel sets A.

In that case, the function mapping μ-measurable sets to ν-measurable sets is well-defined and computable: any μ-approximation of A at precision $2^{-\varphi(n)}$ is a ν-approximation of A at precision 2^{-n}. It implies that every effectively μ-approximable set $A \subseteq X$ or function $f : X \to Y$ is also effectively ν-approximable.

Observe that if $\nu \leq c\mu$ for some constant c then ν is effectively absolutely continuous w.r.t. μ and $\frac{d\nu}{d\mu}$ is essentially bounded by c. The next result is indirectly proved in [87].

Proposition 2.15. *Let μ be a computable probability measure and f be a non-negative $L^1(X,\mu)$-computable. The measure ν defined by $\frac{d\nu}{d\mu} = f$ is computable and effectively absolutely continuous w.r.t. μ.*

Proof. In order to prove that ν is computable, we show that the integral $\int h \,\mathrm{d}\nu = \int fh \,\mathrm{d}\mu$ of any computable bounded function $h : X \to [0,1]$ is uniformly computable. One has $\int f \,\mathrm{d}\mu = \int fh \,\mathrm{d}\mu + \int f(1-h) \,\mathrm{d}\mu$. Both fh and $f(1-h)$ are effectively μ-approximable, so their integrals are lower semicomputable. As their sum is computable, they are computable as well.

We now show that ν is effectively absolutely continuous w.r.t. μ. Given ε, as f is L^1-computable one can effectively find a bounded computable function $h : X \to \mathbb{R}$ such that $\int |f-h| \,\mathrm{d}\mu < \varepsilon/2$. Let N be an upper bound on $|h|$ and $\delta = \varepsilon/(2N)$. If $\mu(A) \leq \delta$ then

$$\nu(A) = \int_A f \,\mathrm{d}\mu = \int_A (f-h) \,\mathrm{d}\mu + \int_A h \,\mathrm{d}\mu \leq \varepsilon/2 + N\mu(A) \leq \varepsilon. \qquad \square$$

However, the Radon-Nikodym theorem is not computable, even assuming effective absolute continuity.

Theorem 2.16 (Hoyrup, Rojas, Weihrauch [70]). *There exists a computable probability measure over* $[0,1]$*, effectively absolutely continuous w.r.t. the Lebesgue measure* λ*, whose Radon-Nikodym derivative is not* $L^1([0,1],\lambda)$*-computable.*

As often, the proof uses the technique appearing in the proof of Pour-El and Richards First Main Theorem [100] stating that a certain class of discontinuous operators do not preserve computability: the discontinuity can be used to encode the halting problem.

Proof. The operator mapping $f \in L^1([0,1],\lambda)$ to the measure ν with density f is continuous (and even computable), but its inverse is not. For instance, the sequence $f_n(x) = 1 + \sin(2\pi nx)$ does not converge to $f_\infty(x) = 1$ in L^1 but the corresponding measures ν_n converge to ν_∞.

Now for $n \in \mathbb{N}$, let $t(n) \in \mathbb{N} \cup \{\infty\}$ be the halting time of Turing machine number n. The sought density function is $f = \sum_{n \geq 1} 2^{-n} f_{t(n)}$. One can easily see that the corresponding measure $\nu = \sum_{n \geq 1} 2^{-n} \nu_{t(n)}$ is computable because the mapping $n \mapsto \nu_{t(n)}$ is computable. However, f is not L^1-computable as $\int |f-1| \, \mathrm{d}\lambda = \frac{2}{\pi} \sum_{n:M_n \text{ halts}} 2^{-n}$ is not a computable real number.

Observe that $|f| \leq 2$, so $\nu \leq 2\lambda$ hence ν is effectively absolutely continuous w.r.t. λ.

Moreover the operator mapping an absolutely continuous measure to its derivative is strongly Weihrauch equivalent to the operator lim.

2.2.5 Effective convergence

Many theorems in measure and probability theory are about convergence of functions or random variables. There are many types of convergence, the most classical ones being:

- Convergence in L^p-norm,
- Almost sure convergence,
- Convergence in probability,
- Convergence in distribution.

In order to carry out a computable analysis of convergence theorems, one has to define effective versions of these notions. Usually a sequence converges to a limit if for every prescribed precision one can find a rank from which the sequence is close to the limit within that precision. This formulation has an immediate effective version, where the rank can be uniformly computed from the given precision. This is the approach we present here. We will see in Section 3 another way of analyzing the computable content of almost sure convergence theorems, using algorithmic randomness. The next definition appeared in [112].

Definition 2.17. Let μ be a probability measure over X and $f_n, f : X \to \mathbb{R}$ be measurable functions. We say that f_n converges ***effectively*** $\boldsymbol{\mu}$***-almost surely*** to f if

$$\forall \varepsilon > 0, \exists n, \quad \mu(\{x \in X : \exists p \geq n, |f_p(x) - f(x)| > \varepsilon\}) \leq \varepsilon \tag{1}$$

and $n \in \mathbb{N}$ can be uniformly computed from $\varepsilon \in \mathbb{Q}$.

We say that f_n converges ***effectively in probability*** to f if

$$\forall \varepsilon > 0, \exists n, \forall p \geq n, \quad \mu(\{x \in X : |f_p(x) - f(x)| > \varepsilon\}) \leq \varepsilon \tag{2}$$

and $n \in \mathbb{N}$ can be uniformly computed from $\varepsilon \in \mathbb{Q}$.

Observe that (1) is equivalent to the more usual formulation of μ-almost sure convergence

$$\mu(\{x \in X : f_n(x) \text{ converge to } f(x)\}) = 1. \tag{3}$$

As in the classical setting, these effective convergence notions are interrelated.

- Effective almost sure convergence implies effective convergence in probability,
- Effective convergence in L^p-norm implies effective convergence in probability,
- When the sequence is bounded by an L^p-computable function, effective convergence in probability implies (therefore, is equivalent to) effective L^p-convergence.

We will see that many convergence theorems from probability theory are not computable, in the sense that the convergence is not effective in any sense (Theorems 2.21, 3.7, 3.8). Such negative results can often be proved by showing that the limit is not effectively approximable, thanks to the following result, appearing in [67].

Proposition 2.18. *Let $f_n : X \to \mathbb{R}$ be uniformly effectively μ-approximable. If they converge effectively in probability to $f : X \to \mathbb{R}$ then f is effectively μ-approximable.*

Proof. To compute f with probability of error δ and at precision ε, compute f_n at precision $\varepsilon/2$ with probability of error δ, where n is associated with $\varepsilon/2$ in the effective convergence.

The proof of this proposition is essentially the argument that we used to show that the limit of the random harmonic series is effectively approximable (Example 2.3).

2.3 Results in computable measure and probability theory

Many results from measure theory have been investigated in computable analysis. Some of them are about the computability of certain measures, others are about the computability of convergence theorems.

The operation of conditioning a measure is a fundamental construct in measure theory. Its (non-)computability has been investigated in [1, 2].

In addition to the results presented here, the computability of invariant measures of dynamical systems has been investigated in [55, 56]. The problem of computing pseudo-random points, i.e. points satisfying prescribed properties that hold almost surely, has been investigated in [12, 52, 54].

2.3.1 Conditioning

Conditioning is a fundamental concept in probability theory. Its (non-)computability has been investigated by Ackerman, Freer and Roy [1, 2]. They prove that the operation of conditioning a probability measure is not computable in general. We present a similar result, whose proof is based on the non-computability of the Radon-Nikodym theorem (Theorem 2.16).

Let π is a Borel probability measure over a product space $X \times Y$ and π_X its marginal measure over X defined by $\pi_X(A) = \pi(A \times Y)$ for Borel sets $A \subseteq X$. One can define the conditional probability measures $\pi(.|x)$ over Y for π_X-almost all $x \in X$, such that π is a combination of these measures: $\pi(A \times B) = \int_A \pi(B|x)\,\mathrm{d}\pi_X(x)$. The function from X to $\mathscr{M}_1(X)$ mapping x to $\pi(.|x)$ is called a ***disintegration*** of π. It is not unique, but two such disintegrations must agree π_X-almost everywhere. Usually these mappings are not continuous, so they cannot be computable. In [1] it is proved that even when a disintegration is discontinuous on a set of measure 1, it need not be computable on a set of measure 1. In [2] it is proved that even when there is a unique continuous disintegration, it need not be computable. The non-computability is also expressed in terms of Weihrauch degrees: the disintegration operator is strongly Weihrauch equivalent to lim.

We present another example, based on the non-computability of the Radon-Nikodym derivative. Indeed, conditional probabilities are usually constructed using the Radon-Nikodym theorem, thus Theorem 2.16 immediately implies the non-computability of conditional probabilities.

Theorem 2.19. *Let $X = [0,1]$ and $Y = \{0,1\}$. There is a computable measure π over $X \times Y$ whose disintegration $x \mapsto \pi(0|x)$ is not effectively π_X-approximable, where π_X is the marginal measure over X (and is the Lebesgue measure λ here).*

Proof. Let ν be the computable probability measure over $[0,1]$ from the proof of Theorem 2.16: $\nu \ll \lambda$ and even $\nu \leq 2\lambda$, but $\frac{d\nu}{d\lambda}$ is not $L^1(\lambda)$-computable. As $\frac{d\nu}{d\lambda}$ is bounded by 2, $\frac{d\nu}{d\lambda}$ is not effectively λ-approximable.

As $\nu \leq 2\lambda$, one has $\lambda = \frac{1}{2}(\nu + \mu)$ where μ is another computable probability measure (μ is simply defined as $2\lambda - \nu$). Consider the measure π over $X \times Y$ defined by $\pi(A \times \{0\}) = \frac{1}{2}\nu(A)$ and $\pi(A \times \{1\}) = \frac{1}{2}\mu(A)$. π is a computable measure, its marginal measures are the uniform measures over X and Y. The conditional expectation is $\pi(0|x) = \frac{d\nu}{d\lambda}(x)$ for λ-almost every $x \in [0,1]$. $\square$

2.3.2 Birkhoff ergodic theorem

One of the most celebrated results in probability theory is the Birkhoff ergodic theorem generalizing the strong law of large numbers from independent random variables to stationary ones.

Ergodic theory is a branch of dynamical systems that focuses on the global properties of dynamical systems (we refer to the introductory book [25]). A (discrete-time) dynamical system is just a set X and a function $T : X \to X$ acting on X.

Points in X are the possible states of the system, and $T(x)$ is the state of the system at time $t+1$ if x is the state at time t. The orbit of a point x is the sequence $x, T(x), T^2(x), \dots$, which is simply the evolution of the system over time when starting in state x. Ergodic theory enables one to describe how the orbits of the system are distributed over X.

To do this, one needs some structure on the set X. Assume that X is a measurable space and μ is a probability measure over X. A measurable transformation $T : X \to X$ ***preserves*** μ if for every Borel set A, $\mu(T^{-1}(A)) = \mu(A)$. We can alternatively say that μ is ***T-invariant***. Intuitively, applying T to points of X does not change their distribution over the space. The Birkhoff ergodic theorem states that if T preserves μ then for μ-almost every x, the asymptotic distribution of the orbit of x under T converges. More precisely,

Theorem 2.20 (Birkhoff ergodic theorem). *If $T : X \to X$ preserves μ then for every $f \in L^1(X, \mu)$, the limit*

$$f^*(x) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(x) \tag{4}$$

exists for μ-almost all $x \in X$. Moreover $f^* \in L^1(\mu)$ and $\int f^* \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu$.

The averages in (4) are called the Birkhoff averages of f.

The case when f is the characteristic function of a measurable set $A \subseteq X$ is particularly suggestive: for μ-almost every x, the visiting frequency of its orbit in A converges, so it is indeed about the distribution of its orbit in the space.

This is an almost-sure convergence theorem. Computing or quantifying the speed of convergence of this theorem has been a longstanding problem, already investigated by ergodic theorists. Kakutani and Petersen [72] proved that there is no general bound on the speed of convergence in the ergodic theorem, but their result does not formally exclude the possibility of a computable speed. They construct very irregular functions f making the convergence as slow as wanted, but it may happen that for simple functions f, the speed of convergence can be estimated. Bishop [19] informally argued that the ergodic theorem is nonconstructive, thus suggesting that it is not computable in any sense. This was made precise by V'yugin [112] who proved that the speed of convergence is indeed not computable in general.

V'yugin's example is based on one of the simplest dynamical systems, the shift operator from the Cantor space $X = \{0,1\}^{\mathbb{N}}$ to itself mapping a sequence $x_0x_1x_2\dots$ to $x_1x_2x_3\dots$, the same sequence with the first bit removed. The function $f : X \to [0,1]$ is very simply defined by $f(x_0x_1x_2\dots) = x_0$.

Theorem 2.21 (V'yugin [112]). *On the Cantor space $X = \{0,1\}^{\mathbb{N}}$, there is a computable shift-invariant probability measure μ such that the convergence of the Birkhoff averages of $f : x \mapsto x_0$ is not effective in any sense (almost sure, in probability, in L^1-norm).*

Proof. We give the proof because it motivates the next discussions. We present a slight variant of V'yugin's original proof. It again follows the argument of Pour-El and Richards First Main Theorem, already met in Theorem 2.16. Informally, the operator mapping μ to $f^* \in L^1(X,\mu)$ is "not continuous".

Given $p \in [0,1]$, consider the measure μ_p over $\{0,1\}^{\mathbb{N}}$ defined as the distribution of the following random infinite binary sequence. Take $x_0 \in \{0,1\}$ uniformly at random. Once x_i has been drawn, let $x_{i+1} = 1 - x_i$ with probability p, and $x_{i+1} = x_i$ with probability $1-p$.

If $p = 0$ then the sequence will contain only 0's with probability $1/2$, and only 1's with probability $1/2$. If $p > 0$ then the sequence will almost surely contain infinitely many blocks of 0's and 1's, very long if p is close to 0.

The discontinuity comes from the fact that f^* has μ_p-almost surely value $\frac{1}{2}$ for $p > 0$ but has μ_p-almost surely values 0 and 1 for $p = 0$.

We can now encode the halting problem using this discontinuity. For $n \in \mathbb{N}$, let $t(n) \in \mathbb{N} \cup \{\infty\}$ be the halting time of Turing machine number n. The sought measure is

$$\mu = \sum_{n \geq 1} 2^{-n} \mu_{2^{-t(n)}}. \tag{5}$$

One can easily see that μ is computable because the mapping $n \mapsto \mu_{2^{-t(n)}}$ is computable. However, f^* is not $L^1(X,\mu)$-computable which implies that the convergence of the Birkhoff averages to $\overline{f}$ is not computable in any sense (see Proposition 2.18). Proving that f^* is not a computable element of $L^1(\mu)$ is easy: otherwise it would be effectively μ-approximable by Proposition 2.13, from which one derives that the singleton $\{0000\ldots\}$ is effectively μ-approximable, which is not possible as $\mu(\{0000\ldots\}) = \sum_{n:M_n \text{ does not halt}} 2^{-n}$ is not a computable real number. $\square$

In this argument one sees that the asymptotic distribution of the orbit of x depends very much on x: the orbits of $\overline{0}$ and $\overline{1}$ are each concentrated in one point while the orbits of other points are densely distributed in the space. The non-computability comes from here, as one cannot computably distinguish between these different behaviors.

The fact that several different distributions are possible is expressed as the system being *non-ergodic*. More precisely, one says that the system (X,μ,T) is ***ergodic*** if the only measurable sets A such that $T^{-1}(A) = A$ are trivial, i.e. have measure 0 or 1. Note that such sets are stable under T, so orbits starting inside A or outside A may have very different behaviors. When the system is ergodic, the limit function f^* is constant μ-almost everywhere, which corresponds to the intuition that almost all the orbits have the same distribution. If (X,T) is fixed, we say that μ is ergodic if the system (X,T,μ) is ergodic.

Let us go back to V'yugin's result. The constructed measure μ is *not* ergodic, notably because the limit function f^* is not constant μ-almost everywhere. This is actually mandatory to make the convergence non-computable, as the next result shows. Here X is a computable metric space, μ a computable Borel probability measure over X and $T : X \to X$ is computable and μ-invariant.

Theorem 2.22 (Avigad, Gerhardy, Towsner [7]). *If the system (X,μ,T) is computable and ergodic and f is a computable element of $L^1(\mu)$ then the speed of convergence of the Birkhoff averages of f is computable.*

They actually prove that in the general (i.e. non-ergodic) case, the speed of convergence is always computable relative to the $L^2(\mu)$-norm of f^*, assuming $f \in L^2(\mu)$. A simpler proof in the ergodic case can be found in [53].

2.3.3 Ergodic decomposition

We saw that the ergodic systems are those systems for which almost all orbits have the same distribution. If a system is not ergodic then two orbits may have very different asymptotic properties. But if one groups together all the points whose orbits have a given distribution then these points form an ergodic subsystem. Indeed a non-ergodic system can always be decomposed into disjoint ergodic subsystems. This is expressed by the following result, which is an application of the Choquet theorem from convex analysis [98].

Theorem 2.23 (Ergodic decomposition theorem). *If μ is a T-invariant probability measure then there exists a unique probability measure m over the class of measures such that:*

- *m gives measure 1 to the set of ergodic T-invariant measures*
- *μ is the barycenter of m, i.e. $\mu(A) = \int \nu(A)\,\mathrm{d}m(\nu)$ for all measurable sets A.*

The ergodic measures can be equivalently defined as the invariant measures that cannot be expressed as combinations of invariant measures, except the trivial one where m is the Dirac measure concentrated on μ.

For instance in V'yugin's construction, the measure μ is a combination of countably many ergodic measures: the Dirac measures $\delta_{000\ldots}$ and $\delta_{111\ldots}$ and for each $t \in \mathbb{N}$, the measure with parameter $p = 2^{-t}$. The decomposition (5) $\sum_{n\geq 1} 2^{-n}\mu_{2^{-t(n)}}$ is computable, but *is not* the ergodic decomposition of μ: for $t(n) = \infty$, one has $\mu_{2^{-t(n)}} = \mu_0 = \frac{1}{2}\delta_{000\ldots} + \frac{1}{2}\delta_{111\ldots}$ which is not ergodic, as it can be further decomposed as a combination of two ergodic measures $\delta_{000\ldots}$ and $\delta_{111\ldots}$. The corresponding ergodic decomposition of μ turns out to be non-computable, and this is necessary and sufficient for the counter-example to work, as the following result shows.

Proposition 2.24 (Hoyrup [64]). *Let T be computable and μ be a computable T-invariant measure. The following statements are equivalent:*

- *The μ-almost sure convergence of the Birkhoff averages of bounded computable functions is effective,*
- *The ergodic decomposition of μ is computable.*

Observe that an invariant measure is usually decomposed into continuously many ergodic measures. In V'yugin's construction, the ergodic decomposition of μ is countably infinite. Is it possible to build another example with a finite number of ergodic measures only? The answer is positive, but one needs a very different argument.

Theorem 2.25 (Hoyrup [65]). *There exists a computable shift-invariant measure μ whose ergodic decomposition is $\mu = \frac{1}{2}\mu_0 + \frac{1}{2}\mu_1$, where μ_0 and μ_1 are non-computable ergodic measures.*

The proof borrows a construction scheme from computability theory, namely the priority method with finite injury.

There are interesting classes of invariant measures for which the ergodic decomposition is computable [62].

Theorem 2.26. *Let $\mathscr{C} \subseteq \mathscr{M}_1(X)$ be an effective compact class of ergodic measures. If μ is a computable measure which is a convex combination of measures in $\mathscr{C}$ then the ergodic decomposition of μ is computable.*

Proof. By the assumption on $\mathscr{C}$, the class of Borel probability measures supported in $\mathscr{M}_1(\mathscr{C})$ is effectively compact in the computable metric space $\mathscr{M}_1(\mathscr{M}_1(X))$ of measures over $\mathscr{M}_1(X)$. The combination operator mapping $m \in \mathscr{M}_1(\mathscr{M}_1(C))$ to $\mu \in \mathscr{M}_1(X)$ is computable and one-to-one, so its inverse is computable. $\square$

Example 2.27 (Pólya urn continued). The class of Bernoulli measures over $\{0,1\}^{\mathbb{N}}$ is an example of an effectively compact class of ergodic measures. In the Pólya urn model from Example 2.4, the measure μ is a combination of Bernoulli measures and is computable, so by Theorem 2.26 its decomposition is computable. As a Bernoulli measure can be (computably) identified with its parameter in $[0,1]$, it means that the distribution of the function $p(s)$ from Example 2.4 is a computable measure over $[0,1]$.

A Bernoulli measure over $\{0,1\}^{\mathbb{N}}$ is the joint distribution of a sequence of i.i.d. random variables in $\{0,1\}$. They can be generalized by considering the distribution of a sequence of i.i.d. random variables in $\mathbb{R}$. These distributions are exactly the product measures $\nu^\infty := \bigotimes_{i\in\mathbb{N}} \nu$, where ν is any probability measure over $\mathbb{R}$. This class of measures $\mathscr{C} = \{\nu^\infty : \nu \in \mathscr{M}_1(\mathbb{R})\} \subseteq \mathscr{M}_1(\mathbb{R}^{\mathbb{N}})$ is no longer effectively compact. However Freer and Roy proved a version of Theorem 2.26 for this class of measures.

Theorem 2.28 (Freer, Roy [47]). *If μ is a computable measure which is a combination of measures in $\mathscr{C}$ then the decomposition of μ is computable.*

De Finetti's theorem states that the convex combinations of measures in $\mathscr{C}$ are exactly the joint distributions of exchangeable sequences of random variables in $\mathbb{R}$, so Theorem 2.28 can be reformulated as follows: if the joint distribution μ of an exchangeable sequence of random variables in $\mathbb{R}$ is computable, then the unique measure m over $\mathscr{C}$ such that $\mu(A) = \int \nu(A)\,\mathrm{d}m(\nu)$ is computable.

3 Algorithmic randomness

Probability theory provides many important convergence theorems, notably almost sure convergence theorems, such as the strong law of large numbers, Birkhoff's ergodic theorem, the martingale convergence theorems, or the Lebesgue differentiation theorem.

In this section, we will work in a probability space, but many of the ideas extend to other types of measures. Recall that a sequence $(f_n)_{n\in\mathbb{N}}$ of random variables converges almost surely to f if $f_n(x)$ converges to $f(x)$ for almost every x. How does one approach the computability of such a result? The option adopted in Definition 2.17 is to use the following equivalent formulation:

$$\forall \varepsilon > 0, \exists n, \quad \mathbb{P}[\forall p \geq n, |f_p(x) - f(x)| \leq \varepsilon] \geq 1 - \varepsilon.$$

This is an existence statement, stating the existence of n given ε, and as such it can be analyzed from a computability perspective. We already saw computability results about almost sure convergence theorems, which happened to be negative for the most part: one cannot compute in general the speed of convergence.

There is another way of formulating almost sure convergence as an existence result:

$$\exists A, \mathbb{P}[A] = 1 \text{ and } \forall x \in A, f_n(x) \text{ converges to } f(x). \tag{6}$$

How does one investigate the computability of such a result? What does it mean to compute the set A? As A has measure 1, it is trivially an effectively measurable set in the sense of Section 2.2.3. Indeed, being effectively measurable is not about a set of points, but about its equivalence class. As a result, one needs a finer effective notion of measurable sets, in particular a notion of effective sets of full measure and effective null sets. This is one of the successes of algorithmic randomness. The reference books in this field are [82, 95, 32], where the theory is developed on the Cantor space. For the extension to more general spaces, we refer the reader to [58, 49, 68, 88, 103, 104].

3.1 Effective null sets

A null set is a set of measure zero. As we are working with regular measures, another way to characterize a null set is to say that a set N is null if and only if for every $\varepsilon > 0$, the set N can be covered by an open set of measure less than ε. Martin-Löf [85] noticed that this characterization of null set can be effectivized.

We are working in any computable metric space X endowed with a computable Borel probability measure μ over X. For short, we say that (X, μ) is a ***computable probability space***. As is usual in probability theory, we sometimes write $\mathbb{P}[A]$ for $\mu(A)$, where $A \subseteq X$ is any Borel set.

Definition 3.1. A ***Martin-Löf test*** is a sequence $(U_n)_{n\in\mathbb{N}}$ of uniformly effectively open sets such that $\mu(U_n) \leq 2^{-n}$. A set N is ***Martin-Löf null*** if $N \subseteq \bigcap_n U_n$ for some Martin-Löf test. A point x is ***Martin-Löf random*** if it is not contained in any Martin-Löf null set. The set of Martin-Löf random points is written ML.

As the name suggests, Martin-Löf random points generally have "typical" or "random" behavior. For example, consider a Martin-Löf random sequence in the space of fair-coin tosses. Such a sequence will always satisfy standard randomness criteria such as the strong law of large numbers and the law of the iterated logarithm. Indeed, it is difficult to come up with an almost sure property of random coin flips which is not true of a Martin-Löf random sequence, and such properties are almost always computability theoretic in nature. (For example, there is a Martin-Löf random which computes the halting problem.) In this definition we have assumed the measure μ to be computable. However it can be extended to non-computable measures by requiring the sets U_n to be effective with oracle μ (we define it below, see Definition 3.16). When the measure μ is not clear from the context, we speak about μ-tests, μ-null sets and μ-random points, whose set is denoted by $\mathsf{ML}(\mu)$.

Nonetheless, for good reasons other randomness notions have also been explored. While there are many options, we give the six that are most connected to computable measure theory. A ***2-Martin-Löf test*** is the same as a Martin-Löf test except that the sequence $(U_n)_{n\in\mathbb{N}}$ is allowed to be computable relative to the halting problem. The corresponding notion of randomness is called 2***-randomness***. A ***weak*** 2***-test*** is a null Π^0_2 set, that is a null set which is the intersection of a sequence of uniformly effectively open sets. The corresponding randomness notion is called ***weak*** 2***-randomness***. Both of these are stronger notions than Martin-Löf randomness (also known as 1-randomness).

The other three notions are weaker. A ***Schnorr test*** is a Martin-Löf test $(U_n)_{n\in\mathbb{N}}$ where $\mu(U_n)$ is computable uniformly in n. The corresponding notion of randomness is called ***Schnorr randomness***. This definition is modeled after Brower's definition of a constructive null set. While weaker than Martin-Löf randomness, Schnorr randomness has a tight relationship with computable measure theory.

Computable randomness is a randomness notion that stands strictly in between Schnorr and Martin-Löf randomness. A test[1] for computable randomness is a Martin-Löf test $(U_n)_{n\in\mathbb{N}}$ which is "bounded" by some computable measure ν. Specifically, $\mu(U_n \cap A) \leq \nu(A) \cdot 2^{-n}$ for all Borel measurable sets A.

Last, ***Kurtz randomness*** (or ***weak randomness***) is the weakest randomness notion of the six. A Σ^0_2 set is a computable union of effectively closed sets. A ***Kurtz test*** is a null Σ^0_2 set. A set is ***effectively Kurtz null*** if it is a subset of a Kurtz test. A point is ***Kurtz random*** if it is not contained in an effectively Kurtz null set. Kurtz randomness is much weaker than Schnorr randomness, and many do not consider it to be a true randomness notion. It does not satisfy the law of large numbers, and it

[1] On Cantor space, tests are usually expressed in terms of computable martingales. On general metric spaces, the equivalent definition given here is easier to express.

shares as many similarities with effectively genericity as it does with randomness.[2] Nonetheless, it is useful to consider Kurtz randomness.

To summarize we have the following algorithmic randomness notions listed from weakest to strongest: Kurtz, Schnorr, computable, Martin-Löf, weak 2-, and 2-randomness. While the majority of the work in algorithmic randomness has focused on Cantor space with the fair-coin (a.k.a. Lebesgue) measure (or in some cases Bernoulli measures), all of these randomness notions naturally extend to other computable metric spaces endowed with a Borel probability measure. One natural candidate is Brownian motion.

Brownian motion.

Informally Brownian motion is a process which resembles a continuous time random walk starting at the origin. More formally one can represent a d-dimensional Brownian motion as a particular probability measure, called the Wiener measure, on the space $\mathscr{C}([0,1],\mathbb{R}^d)$ (or $\mathscr{C}([0,\infty),\mathbb{R}^d)$) of continuous functions. This space is a computable metric space under the sup norm. Also the Wiener measure is a computable probability measure on this space. The algorithmically random Brownian motion paths have been thoroughly studied by Fouché and others [4, 36, 37, 77, 38, 39, 28, 13, 40, 42, 41].

To give a sense of the techniques involved let us consider recurrence and transience of Brownian motion in 2 and 3 dimensions. (This theorem is sometimes stated as "A drunk person will always find their way home, while a drunk bird never will.")

Theorem 3.2. *Let $B\colon [0,\infty) \to \mathbb{R}^d$ be a continuous function.*

1. *If $d = 2$ and B is Kurtz random, then B obeys the following recurrence result: For every ε and every $t_0 \geq 0$, there is some $t > t_0$ such that $|B(t)| < \varepsilon$.*
2. *If $d \geq 3$ and B is Schnorr random, then B obeys the following transience result: $|B(t)| \to \infty$ as $t \to \infty$.*

Proof. In both cases, our goal is to analyze the corresponding null set. Also, to avoid "reinventing the wheel" we will use known results in Brownian motion, including the classical theorems we are attempting to effectivize.

For (1), consider a non-recurrent function B in two dimensions. Then there are natural numbers m and n such that $|B(t)| \geq 2^{-n}$ for all $t \geq m$. By standard techniques

[2] This claim that Kurtz randomness is not really a randomness notion can be made formal by considering the relativized versions of the tests. Every null set is a Martin-Löf null set relative to some oracle. Said another way, the only difference between the definition of a null set and a Martin-Löf null is computability. The same is true for 2-randomness, weak 2-randomness, computable randomness, and Schnorr randomness. However, this is not true for Kurtz randomness. The relativized notion of a Kurtz null set is a null F_σ set. It is well known in measure theory that there are null sets which are not contained in any null F_σ set. Indeed, every null F_σ set is meager which explains the connections with genericity. There still is a connection with analysis. Null F_σ sets are the type of null set associated with Jordan-Peano measurable sets and Riemann integrable functions, as opposed to Lebesgue measurable sets and Lebesgue integrable functions.

in computable analysis this is a Σ_2^0 property, i.e. its complement is the intersection of a sequence of uniformly effective open sets. By the recurrence theorem on 2-dimensional Brownian motion, the set of such paths is null. Therefore, we have the desired Kurtz null test.

For (2), consider a non-transient function B in three dimensions. If B is Kurtz random, then by the same argument as in the recurrent case, $|B(t)|$ is unbounded. Therefore, for every $0 < r < R$, there is a $t_0 > 0$ such that $|B(t_0)| = R$ and a $t_1 > t_0$ such that $|B(t_1)| = r$. Now we will use the following quantitative estimate from probability theory:

$$\mathbb{P}\underbrace{\{B : \exists t_1 > t_0 > 0\ [|B(t_0)| > R \text{ and } |B(t_1)| < r]\}}_{U_{r,R}} \leq \frac{r^{d-2}}{R^{d-2}}$$

A quick observation is that $U_{r,R}$ is effectively open uniformly in r and R (which we can assume are rationals). Then $\bigcap_{r,R} U_{r,R}$ is a null Π_2^0 set. Therefore, every weak 2-random Brownian motion path is transient. We can do better. Since we have a computable upper bound on $\mathbb{P}[U_{r,R}]$ by picking a sufficiently fast shinking/growing set of pairs $r_n \to 0$ and $R_n \to \infty$, we have a Martin-Löf test $U_n = U_{r_n,R_n}$.

This is the point where most of the results in the literature stop. However, with a little more work we can extend our result to Schnorr randomness by showing that $\mathbb{P}[U_{r,R}]$ is computable in r and R. A common misconception is that we require an *exact formula* here. Indeed, we only require an *algorithm* which converges to the value of $\mathbb{P}[U_{r,R}]$. We provide such an algorithm. Since $U_{r,R}$ is effectively open, $\mathbb{P}[U_{r,R}]$ is lower semicomputable. Therefore it remains to show $\mathbb{P}[U_{r,R}]$ is upper semicomputable. Choose a large $S > R$ and a small $\varepsilon > 0$. By transience, after $|B(t)| = R$, then almost surely, eventually there is some $t_1 > t$ such that $|B(t_1)| > S$. Let $V_{r,R,S}$ be the set of paths which after hitting R, they hit S before r. This set is also effectively open, and therefore $\mathbb{P}[V_{r,R,S}]$ is lower semicomputable.[3] Now, almost every path is in either $U_{r,R}$ or $V_{r,R,S}$ (or both!). However, the overlap of the two sets is small since $U_{r,R} \cap V_{r,R,S} \subset U_{r,S}$. By choosing S large enough we can compute $\mathbb{P}[U_{r,R}]$ to any desired precision. □

This previous proof is typical for this type of result. Most of the results for Martin-Löf randomness in the literature of this type can be strengthened to Schnorr randomness with some additional work.[4]

---

[3] Formally $V_{r,R,S} = \{B : \exists t_0 < t_1\ [r < |B(t_0)| < R \quad \wedge \quad S < B(t_1) \quad \wedge \quad \min_{t\in[t_0,t_1]} |B(t)| > r]\}$. This is effectively open since we are using $<$ signs and since the the minimum of a continuous function over a closed interval is computable. Also, if we replaced any of the $<$ with $\leq$ it would not change the measure of the set in this case.

[4] By "this type of result", we mean an almost everywhere result for some computable probability space *which does not explicitly refer to computable objects*. Effective convergence results (see the next section) explicitly refer to computable objects and may hold for Martin-Löf randomness but not Schnorr randomness.

3.2 Effective convergence theorems

As explained at the beginning of Section 3, algorithmic randomness provides a way to analyze the computability of almost sure convergence theorems, by identifying the notion of effective null set associated with each such theorem. There has been increasing interest in the past few years on obtaining characterizations of almost sure convergence theorems using notions of algorithmic randomness. We present here some of the most prominent results in the literature, characterizing Martin-Löf, computable and Schnorr randomness using Birkhoff's ergodic theorem and differentiation theorems, but there are other works calibrating the notions of randomness to fit almost sure theorems with varying computability assumptions.

The first results in this direction have been obtained very early by characterizing algorithmic randomness notions in terms of convergence of martingales. On the Cantor space, Schnorr has proved that the Martin-Löf random points are exactly the points where every *lower semicomputable* martingale is bounded. He introduced computable random points as the points at which every *computable* martingale is bounded. He also introduced what is now called Schnorr randomness, also originally expressed in terms of martingales. We do not give the details of these characterizations, which can be found in the reference books [82], [95] or [32]. Schnorr's work on this topic appeared in his book [106].

3.2.1 Ergodic theorems

As we saw in Section 2.3.2, Birkhoff's ergodic theorem is not computable in general, i.e. the convergence is not effective in any sense. What about the effectiveness of the associated null set? In other words, for which class of random points does Birkhoff's ergodic theorem hold?

Bishop already observed that this theorem does not hold constructively. He proposed a constructive proof, whose conclusion is classically equivalent to almost sure convergence, but has less constructive content. This statement involves so-called upcrossing inequalities. This idea was later used by V'yugin to prove that the Birkhoff ergodic theorem holds for Martin-Löf random points.

Theorem 3.3 (V'yugin [112]). *Let X be a computable metric space, $T : X \to X$ a computable function and μ a T-preserving probability measure over X and $f : X \to \mathbb{R}$ a computable function. For every Martin-Löf μ-random point x, the limit*

$$f^*(x) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(x)$$

exists.

In other words, the associated null set is a Martin-Löf null set. There has been work investigating the classes of functions f to which Theorem 3.3 can be extended. It holds for all bounded continuous (not necessarily computable) functions f, just because computable functions are in a sense dense among them. It was proved in [14, 43, 15] that it also holds when the system is ergodic and $f : X \to [0,+\infty]$ is lower semicomputable.

The result was used in [63] to prove another famous result from ergodic theory, the Shannon-McMillan-Breiman theorem, for Martin-Löf random points. That result was already implicitly proved in [60] using upcrossing inequalities. We will see in Section 4.5 that Theorem 3.3 can be extended to other classes of functions T, f that are not continuous (layerwise computable functions). It is still open whether that result holds for lower semicomputable $f : X \to [0,+\infty]$ when the system is not ergodic, and for upper semicomputable $f : X \to [0,+\infty]$ (both for ergodic and non-ergodic systems). It is proved in [89] that this result holds for a stronger notion of randomness called Oberwolfach randomness.

A converse to Theorem 3.3 was later proved by Franklin and Towsner [44], showing that Martin-Löf randomness is the right notion in that case.

Theorem 3.4 (Franklin, Towsner [44]). *If x is not Martin-Löf λ-random then there exists a computable λ-preserving map $T : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ and an effective open set A such that $\lambda(A)$ is computable and such that the Birkhoff averages of $\mathbf{1}_A$ do not converge at x.*

In other words, if a class of effective null sets induces a notion of randomness that is not stronger than Martin-Löf randomness (if for instance it is strictly weaker than Martin-Löf randomness, like Schnorr or computable randomness), then the Birkhoff ergodic theorem is not in general effective for this notion of null set.

We saw that in the ergodic case, the speed of convergence is computable. It has a consequence on the corresponding randomness notion.

Theorem 3.5 (Gács, Hoyrup, Rojas [50]). *Let X be a computable metric space and μ a computable probability measure over X. A point x is Schnorr μ-random if and only if for every μ-preserving computable $T : X \to X$ and every bounded continuous $f : X \to \mathbb{R}$ the limit*

$$f^*(x) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(x)$$

exists.

In other words, if a class of effective null sets induces a notion of randomness that is not stronger than Schnorr randomness then the Birkhoff ergodic theorem for ergodic measures is not in general effective for this notion of null set.

3.2.2 Differentiation theorems

In this section the underlying measure is the Lebesgue or uniform measure over the real interval $[0,1]$. Let us start with a computable theorem, the Lebesgue differentiation theorem.

Theorem 3.6 (Pathak, Rojas, Simpson [97]; Freer, Kjos-Hanssen, Nies, Stephan [45]). *For a real $x \in [0,1]$ the following statements are equivalent:*

- *x is Schnorr random,*
- *For every L^1-computable function $f : [0,1] \to \mathbb{R}$,*

$$\lim_{r \to 0} \frac{1}{\lambda(B(x,r))} \int_{B(x,r)} f \, \mathrm{d}\lambda \text{ exists.}$$

As already mentioned, when considering almost sure convergence theorems one quickly runs into non-computability results. The following theorems are not computable in the sense that the almost sure convergence is not effective. How to distinguish between different types of non-computability? Again, algorithmic randomness allows for a finer look by providing several notions of randomness via several notions of effective null sets.

Theorem 3.7 (Brattka, Miller, Nies [24]; Freer, Kjos-Hanssen, Nies, Stephan [45]). *For a real $x \in [0,1]$ the following statements are equivalent:*

- *x is computably random,*
- *Every non-decreasing computable function $f : [0,1] \to \mathbb{R}$ is differentiable at x.*
- *Every computable Lipschitz function $f : [0,1] \to \mathbb{R}$ is differentiable at x.*

Nies [96] proved that this theorem also holds for polynomial randomness and polynomial-time computable functions f.

Theorem 3.8 (Demuth [29]; Brattka, Miller, Nies [24]). *For a real $x \in [0,1]$ the following statements are equivalent:*

- *x is Martin-Löf random,*
- *Every computable function $f : [0,1] \to \mathbb{R}$ of bounded variation is differentiable at x.*

This result was obtained by Demuth [29] in the context of constructive analysis and was reformulated in [24] in modern language.

These results highlight the non-computability of some theorems from real analysis. For instance the Jordan decomposition states that every function of bounded variation is a difference of two non-decreasing functions. The two previous theorems witness that this decomposition is not computable in general. These results are also indirect proofs that the convergence in these theorems is not effective, otherwise convergence would occur at every Schnorr random point.

3.3 Randomness preservation

While "almost sure" results are an important category of results in measure theory, there are also many results which relate the behavior of random variables (or measurable sets) on one space to random variables (or measurable sets) on another space. For example, in probability theory, a random variable representing a sequence of independent fair coin flips can be easily transformed into a random variable representing a random walk on a lattice or a normally distributed random variable. In the context of algorithmic randomness, one wants to be sure that if x is random on one space, then this transformation of x remains random in the corresponding space. From the perspective of computable analysis, we are saying that these transformations preserve effective null sets. Here we give examples of such results.

Let X and Y be computable metric spaces and μ a computable Borel probability measure over X. We recall from Section 2.2.2 that a function $f : X \to Y$ induces a measure μ_f over Y called the push-forward measure, defined by $\mu_f(A) = \mu(f^{-1}(A))$ for all Borel sets $A \subseteq Y$. The next results are folklore results.

Theorem 3.9 (Randomness preservation). *Let $f : X \to Y$ be a computable function. If x is Martin-Löf μ-random, then $f(x)$ is Martin-Löf μ_f-random.*

This result is well-known and holds for Kurtz, Schnorr, Martin-Löf, weak 2-, and 2-randomness. It, however, does not hold for computable randomness (see Rute [103]). There is also a partial converse to randomness preservation.

Theorem 3.10 (No-randomness-from-nothing). *Let $f : X \to Y$ be a computable function. If y is Martin-Löf μ_f-random, then there exists Martin-Löf μ-random x such that $f(x) = y$.*

No-randomness-from-nothing holds for computable, Martin-Löf, weak 2-, and 2-randomness, but not for Schnorr randomness (see Rute [103]). Nonetheless, we will see a result below (Theorem 3.18) that implies in most natural cases that no-randomness-from-nothing holds for Schnorr randomness.

Observe that Proposition 2.9 implies that the measure μ_f is computable. Proposition 2.9 holds not only for computable functions, but for the larger class of effectively μ-approximable functions. However, Theorems 3.9 and 3.10 cannot hold for those functions: if f is in that class and x is Martin-Löf μ-random then $f(x)$ can be anything, unless $\mu(\{x\}) > 0$ (f remains in the class when changing it at x only). Here we see that one needs a notion of effectively measurable function that is well-behaved on algorithmically random points. Such a notion exists and is called *layerwise computability*, presented in Section 4.5.

To give a specific application of these results, let us introduce the notion of a random closed set.

Example 3.11 (Random closed sets). In the context of computability theory, Barmpalias, Broadhead, Cenzer, Dashti, and Weber [10] introduced the following definition of a "(Martin-Löf) random closed set". Every infinite tree T with no dead ends on $\{0,1\}^*$ can be represented by a ternary sequence X in $\{0,1,2\}^{\mathbb{N}}$ recursively as follows. Start at the root of T. If at the root T branches to both the left and the right, let $X(0) = 0$. It the only branches to the left, $X(0) = 1$ and if it only branches to the right, $X(0) = 2$. Recursively, for each node of the tree, similarly define a value of X to code how the tree branches at that node. If the sequence X is (Martin-Löf) random, then the set $C \subseteq \{0,1\}^{\mathbb{N}}$ of paths through this tree is called a ***(Martin-Löf) random closed set***.

The space $\mathscr{K}(2^{\mathbb{N}})$ of non-empty closed subsets of $2^{\mathbb{N}}$ forms a computable metric space with the Hausdorff metric (the topology is called the Fell topology). Moreover, the map $X \mapsto C$ which maps the representation for the tree T to its set of paths is computable as a map from $3^{\mathbb{N}}$ to $\mathscr{K}(2^{\mathbb{N}})$. If μ is the push-forward measure along this map, then our above theorems tell us that a closed set C is Martin-Löf random (in the specific sense that its encoding X is random) if and only if C is Martin-Löf μ-random. Actually, since this map is computably invertible, this result will hold for all of the major randomness notions listed in Section 3.1.

While this last example may seem simple, results such as these are key to working with random structures in mathematics. In many cases, they can turn a long proof into a short one. Here is an example.

Example 3.12 (Brownian motion). Consider a one dimensional Martin-Löf random Brownian motion $B : [0,1] \to \mathbb{R}$. It is well-known that the push-forward of the map $B \mapsto B(1)$ induces the Gaussian measure on $\mathbb{R}$. Therefore, it is natural to suspect that for every random Brownian motion B, the value of $B(1)$ is random (for the Gaussian measure) and if a is (Gaussian) random then there is a random Brownian motion such that $B(1) = a$. Indeed, for Martin-Löf randomness, the first fact follows from randomness preservation and the second from no-randomness-from-nothing. Observe that the Gaussian measure and the uniform measure on $\mathbb{R}$ have the same Martin-Löf random points (we will see why in Example 4.7).

We will see in Section 4.5 other randomness preservation results where Theorems 3.9 and 3.10 cannot be applied because the function involved is not continuous, but only measurable.

3.4 Product spaces

It is well known that if one takes two *independent* sequences of independent identically distributed (i.i.d.) fair-coin flips $A = (a_0, a_1, a_2, \ldots)$ and $B = (b_0, b_1, b_2, \ldots)$ and interleave them, then the resulting sequence $A \oplus B = (a_0, b_0, a_1, b_1, \ldots)$ is also an i.i.d. sequence of fair-coin flips. The concept of independence is essential to probability theory, and its analogue in algorithmic randomness is relative randomness.

Definition 3.13 (Relative Martin-Löf randomness). Let (X, μ) be a computable probability space and Y a computable metric space. A ***uniform Martin-Löf test*** is a computable sequence of effectively open sets $(U_n)_{n\in\mathbb{N}}$ on the product space $X \times Y$ such that for every $y \in Y$, we have $\mu\{x \in X : (x,y) \in U_n\} \leq 2^{-n}$. Say that $x \in X$ is ***Martin-Löf random relative to*** y if (x,y) not contained in $\bigcap_n U_n$.

Theorem 3.14 (Van Lambalgen [80]). *Endow* $\{0,1\}^{\mathbb{N}}$ *with the fair-coin measure. For A and B in* $\{0,1\}^{\mathbb{N}}$ *the following are equivalent.*

1. $A \oplus B$ is Martin-Löf random,
2. A is Martin-Löf random and B is Martin-Löf random relative to A.

This result also extends to any computable product measure. Given two probability spaces (X,μ) and (Y,ν), the product measure $\mu \otimes \nu$ is the probability measure on $X \times Y$ given by $(\mu \otimes \nu)(A \times B) = \mu(A) \cdot \nu(B)$. The product measure operation is computable and we have the following version of van Lambalgen's theorem.

Theorem 3.15 (Van Lambalgen for product measures). *Let (X,μ) and (Y,ν) be two computable probability spaces. Let $x \in X$ and $y \in Y$. The following are equivalent.*

1. (x,y) is Martin-Löf $(\mu \otimes \nu)$-random,
2. x is Martin-Löf μ-random and y is Martin-Löf ν-random relative to x.

Last, one may consider the case of an arbitrary probability measure on a product space. We saw in Section 2.3.1 how a probability measure π over $X \times Y$ can be decomposed into its marginal measure π_X over X, defined by $\pi_X(A) = \pi(A \times Y)$, and conditional measures $\pi(\cdots \mid x)$. We saw that the mapping $x \mapsto \pi(\cdot \mid x)$ is not always computable or effectively π_X-approximable. However when it is computable one can prove a version of van Lambalgen's theorem for conditional probabilities.

Definition 3.16. Let X be a computable metric space. A ***uniform Martin-Löf test*** is a computable sequence of effectively open sets $(U_n)_n \in \mathbb{N}$ on the product space $\mathscr{M}_1(X) \times X$ such that for every $\mu \in \mathscr{M}_1(X)$, we have $\mu\{x \in X : (\mu,x) \in U_n\} \leq 2^{-n}$. Say that $x \in X$ is ***Martin-Löf μ-random*** if (μ,x) is not contained in $\bigcap_n U_n$.

One can naturally combine Definitions 3.13 and 3.16 to define Martin-Löf randomness for a noncomputable measure relative to a noncomputable oracle. A version of van Lambalgen's theorem for conditional probabilities was proved by Takahashi [109] on the Cantor space and holds on all computable metric spaces.

Theorem 3.17 (Van Lambalgen's theorem for conditional probabilities). *Let X,Y be computable metric spaces and π a computable measure over $X \times Y$. Assume that $x \mapsto \pi(\cdot \mid x)$ is a computable function. The following are equivalent.*

1. (x,y) is Martin-Löf π-random,
2. x is Martin-Löf π_X-random and y is Martin-Löf $\pi(\cdot \mid x)$-random relative to x.

In this result, the measure $\pi(\cdot \mid x)$ is computable relative to x, *for all $x \in X$ and uniformly in x*. Takahashi [110] extended this theorem by proving that the equivalence holds for a single x as long as $\pi(\cdot \mid x)$ is computable relative to x (while $\pi(\cdot \mid x')$ may not be computable relative to x' for $x' \neq x$), still assuming that the measure π is computable. Bauwens [11] showed that in general the equivalence in Theorem 3.17 fails when $\pi(\cdot \mid x)$ is not computable relative to x.

Now, let us turn to maps and random variables. Let (X,μ) be a computable probability space and Y a computable metric space. Let $T : X \to Y$ be a computable map. The pushforward measure μ_T is computable. Let $\mu(\cdot \mid T = y)$ denote the conditional probability of x given that $T(x) = y$. Even if $\mu_T(y) = 0$ for all y, the conditional probability is well defined as a measurable function $y \mapsto \mu(\cdot \mid T = y)$. In particular, $y \mapsto \mu(\cdot \mid T = y)$ is the μ_T-almost everywhere unique function satisfying

$$\mu(A \cap T^{-1}(B)) = \int_B \mu(A \mid T = y)\, \mathrm{d}\mu_T(y)$$

for all measurable sets $A \subseteq X$ and $B \subseteq Y$.

Theorem 3.18 (Van Lambalgen's theorem for maps). *Let $T\colon (X,\mu) \to Y$ be as above and assume the conditional probability map $y \mapsto \mu(\cdot \mid T = y)$ is computable. Then the following are equivalent.*

1. *x is Martin-Löf μ-random,*
2. *$y := T(x)$ is Martin-Löf μ_T-random and x is Martin-Löf random w.r.t. $\mu(\cdot \mid T = y)$, relative to y.*

Proof. We simply apply the previous theorem to the measure π over $X \times Y$ defined as the push-forward of μ along the map $x \mapsto (x, T(x))$. That is π is the measure supported on the graph of T whose marginal measure is μ. Conditioning π on x is given by $\pi(\cdot \mid x) = \delta_{T(x)}$ and is computable, so (x,y) is π-random iff condition 1. is satisfied (as $T(x)$ is always $\delta_{T(x)}$-random). Conditioning π on y is given by $\pi(\cdot \mid y) = \mu(\cdot \mid T = y)$ and is computable by assumption, so (x,y) is π-random iff condition 2. is satisfied. □

The above mentioned results can be extended in two natural ways. First, all of these versions of Van Lambalgen's theorem hold for Schnorr randomness under the correct notion of "relative Schnorr randomness". See Miyabe and Rute [90] and Rute [104] for details. Second, any theorem requiring that a map is "computable" can be extended to layerwise computable map (or in the case of Schnorr randomness, a Schnorr layerwise computable map), see Definition 4.3. While the details are more technical, this allows for a much more natural setting.

Example 3.19 (Brownian motion again). We have seen in Examples 3.12 that the values $B(1)$ taken by Martin-Löf random Brownian paths B at time 1 are exactly the Martin-Löf random reals w.r.t. the Lebesgue measure. The same result holds when replacing Martin-Löf randomness with Schnorr randomness on both sides. Indeed, the conditional probability of the map $B \mapsto B(1)$ conditioned on $B(1) = a$ is known as a ***Brownian bridge landing at*** a. Such objects are well studied in probability theory and the conditional probability map is computable. Therefore, for Schnorr randomness, the result follows from Van Lambalgen's theorem for maps (see Rute [104] for the details of this result.)

4 Pointwise computable measure theory

Let us take a step back regarding all the material developed so far. Computable measure theory is somewhat separated from other branches of computable analysis in several aspects:

- Computable analysis can usually be seen as computable topology. Indeed, the core concept of computable function in computable analysis is very close to the notion of continuous function, therefore it is appropriate on topological spaces but not on measure spaces.
- In computable analysis, the fundamental notion of computable function is defined in terms of a function mapping names of points to names of their images. However the convenient notion of effectively measurable function does not comply with this definition.
- Computable analysis is generally about computing points or functions between points, however individual points are completely ignored in the definitions of effectively approximable and effectively measurable sets and functions.

We also saw that there is no unique way of investigating convergence theorems in terms of computability and that computable measure theory and algorithmic randomness give quite different insights on this problem. So it seems that we need to reconcile computable measure theory with computable topology and to investigate more precisely the relationship between the two parts of this chapter, computable measure theory and algorithmic randomness. Let us draw inspiration from the following text written by Doob in his book *Measure Theory* ([31] p. 101), comparing analysis before and after the advent of measure theory:

> In many contexts, measure theory widened the class of admissible domains and functions to the classes of measurable sets and measurable functions, and in so doing made it possible to apply the usual limiting procedures without leaving admissible classes. What was unexpected was that, in a reasonable sense, most of the old concepts were very nearly still present. Egorov's theorem showed that uniform convergence was nearly present whenever there was convergence. Lusin's theorem showed that the new measurable functions were nearly continuous. On the other hand, measure theory could be applied in abstract contexts where topology was inappropriate.

This phenomenon has consequences in computable measure theory. We will see that the "old concepts" from computable analysis can be used in computable measure theory, thanks to algorithmic randomness and more particularly Martin-Löf randomness.

This is possible because algorithmic randomness is inherently a pointwise approach to computable probability theory. We will see that many computable versions of results and constructs in measure and probability theory have a pointwise formulation. We have already seen that independence of random variables has a pointwise formulation, namely Van Lambalgen's theorem. We will see that effective absolute continuity of measures has a formulation in terms of randomness preservation, effective Egorov's theorem can be formulated in terms of uniform convergence on random points, effective Lusin's theorem can be expressed as a form of uniform computability on random points, called layerwise computability, and so on.

The results in this section heavily rely on the important notion of randomness deficiency.

Randomness deficiency.

One of the first main results about Martin-Löf randomness is the existence of a universal Martin-Löf test, or a greatest Matin-Löf null set. More precisely there exists a Martin-Löf test $(U_n)_{n\in\mathbb{N}}$ such that for every Martin-Löf test $(V_n)_{n\in\mathbb{N}}$ there exists $c\in\mathbb{N}$ such that $V_{n+c}\subseteq U_n$ for all n. We fix such a universal test and define $\mathsf{ML}_n = X\setminus U_n$. The set of Martin-Löf random points can then be decomposed into levels $\mathsf{ML}=\bigcup_n \mathsf{ML}_n$ with $\mathsf{ML}_n\subseteq\mathsf{ML}_{n+1}$. If x is Martin-Löf random then the minimal n such that $x\in\mathsf{ML}_n$ is called the ***randomness deficiency*** of x. There are actually many equivalent ways of tests for Martin-Löf randomness. In each case there exists a universal test and its associated randomness deficiency notion. While these different quantities are not equal, they are computably related and all the results in this section remain true for these other notions.

The levels have large measures: $\mu(\mathsf{ML}_n)\geq 1-2^{-n}$ and are effective closed sets, or Π^0_1-sets. One can prove more, as we now show.

4.1 Effective tightness

In a complete separable metric space, every Borel probability measure is tight, i.e. assigns most of the weight to compact sets. This theorem is effective and witnessed by the Martin-Löf random points and their deficiencies.

Proposition 4.1 (Effective tightness). *Let X be a complete computable metric space and μ a computable Borel probability measure over X. The sets $\mathsf{ML}_n(\mu)$ are effectively compact, uniformly in n.*

This result appeared in [67]. More generally, Martin-Löf random points witness an effective version of Prokhorov's theorem: if $\mathscr{C}\subseteq\mathscr{M}_1(X)$ is an effectively compact class of Borel probability measures then $\mathsf{ML}_n(\mathscr{C}):=\bigcup_{\mu\in\mathscr{C}}\mathsf{ML}_n(\mu)$ is effectively compact [16]. For this, one needs to define Martin-Löf randomness for non-computable measures (see Definition 3.16 below and [49, 68]).

Proposition 4.1 is fundamental as it enables to apply "old concepts" from computable analysis, involving effective compactness, to computable measure theory. We will see for instance how the effectiveness of the Pólya urn (Example 2.4) can be easily proved by a compactness argument rather than by probabilistic estimates (see Example 4.14).

This result is useful to study Martin-Löf randomness of Brownian motion, where the underlying space $\mathscr{C}([0,1],\mathbb{R}^d)$ is not compact, but the levels of Martin-Löf random paths are effectively compact. For instance, it implies that for any computable Borel probability measure over $\mathscr{C}([0,1],\mathbb{R})$, the Martin-Löf random functions all have a computable modulus of uniform continuity. For the particular case of the Wiener measure, an explicit formula is given by Lévy's modulus of continuity theorem, but the computability of the modulus holds for any computable measure.

4.2 Effective Egorov theorem

If a point x is Martin-Löf random then it has a finite randomness deficiency, which cannot usually be computed or even bounded from an access to x. Moreover, having an upper bound on this deficiency gives important information about x that usually cannot be recovered from x. For instance, Davie [27] showed how this additional information can be used to compute the speed of convergence in the strong law of large numbers or the law of the iterated logarithm, or to bound the number of events in the Borel-Cantelli lemma. More generally one can prove an effective version of Egorov's theorem, involving again Martin-Löf random points and their deficiencies [67].

Proposition 4.2 (Effective Egorov's theorem). *Let $f_n, f : X \to \mathbb{R}$ be uniformly computable functions. The following statements are equivalent:*

1. *f_n converge effectively μ-almost surely to f,*
2. *f_n converge effectively uniformly to f on each $\mathsf{ML}_k(\mu)$, uniformly in k.*

The second item means that given $k, \varepsilon > 0$, one can compute n such that

$$\sup_{x \in \mathsf{ML}_k(\mu)} |f_p(x) - f(x)| < \varepsilon$$

for all $p \geq n$.

Incidentally, this result relates more precisely the effective convergence notions investigated in the two parts of this chapter: effective almost sure convergence (Definition 2.17) and convergence on algorithmically random points. We saw that the former implies the latter (even for Schnorr random points) but the converse does not usually hold, and Proposition 4.2 shows the precise relationship.

4.3 Effective Lusin's theorem

We have just seen in Proposition 4.2 that if f_n are computable and converge effectively almost surely to f then for an individual x, one can compute the speed of convergence of $f_n(x)$ to $f(x)$ simply from any upper bound on its randomness deficiency. In particular one can compute $f(x)$ from x and such an upper bound. We make this observation into a definition [68].

Let (Y, δ_Y) be a set with a representation.

Definition 4.3. A function $f : X \to Y$ is ***μ-layerwise computable*** if for all $k \in \mathbb{N}$ and all $x \in \mathsf{ML}_k(\mu)$, $f(x)$ is computable relative to x, uniformly in k, x.

This means that there is a type-two Turing machine taking as inputs $k \in \mathbb{N}$ and a name of $x \in \mathsf{ML}_k(\mu)$ and outputs a name of $f(x)$.

The next result [66] shows that this notion is an alternative to the notions of effectively approximable and effectively measurable functions. It is at the same time an effective version of Lusin's theorem, which states that every measurable function is *nearly continuous*, i.e. continuous on a compact subset of measure arbitrarily close to 1.

Proposition 4.4 (Effective Lusin theorem). *Let $f : X \to Y$. The following statements are equivalent:*

1. *f is effectively μ-approximable,*
2. *f coincides μ-almost everywhere with a μ-layerwise computable function.*

Note that a μ-layerwise computable function is computable hence continuous on each $\mathsf{ML}_k(\mu)$, so this proposition is indeed an effective version of Lusin's theorem.

There are good reasons to use the notion of layerwise computable function in place of effectively measurable/approximable function. On the one hand it is much more adapted to the study of Martin-Löf randomness. For instance two layerwise computable functions that coincide almost everywhere actually coincide on the Martin-Löf random points. Almost all the theorems about Martin-Löf random points involving computable functions actually hold for layerwise computable functions, thus this notion is the suitable notion of effectively measurable function that is well-behaved w.r.t. Martin-Löf randomness. On the other hand it enables one to use "old concepts" from computable analysis in computable measure theory, because they are instances of computable functions between represented spaces for a suitable representation (a random point is represented by a name and an upper bound on its randomness deficiency), and they are almost like computable functions, with some non-uniformity.

4.4 Effective absolute continuity

We have already seen randomness preservation theorems, which are theorems about preservation of effective null sets. The question of the preservation of null sets is naturally raised in another situation, when comparing two measures. We recall that a measure ν is absolutely continuous w.r.t. a measure μ, written $\nu \ll \mu$, if every μ-null set is also ν-null. It is natural to investigate effective versions, where null sets are replaced by any notion of *effective* null sets.

Theorem 4.5 (Bienvenu, Merkle [17]). *The following implications hold and are tight:*

$$\left.\begin{array}{r}\mathsf{CR}(\nu)\subseteq\mathsf{CR}(\mu)\implies\mathsf{ML}(\nu)\subseteq\mathsf{ML}(\mu)\\ \mathsf{Sch}(\nu)\subseteq\mathsf{Sch}(\mu)\end{array}\right\}\implies\nu\ll\mu\implies\mathsf{Ku}(\nu)\subseteq\mathsf{Ku}(\mu).$$

The implication $\mathsf{ML}(\nu)\subseteq\mathsf{ML}(\mu)\implies\nu\ll\mu$ was independently proved by Archibald, Brattka and Heuberger [3].

We have seen another effective version of absolute continuity in Definition 2.14. It is related to algorithmic randomness through a characterization that involves randomness deficiency.

Proposition 4.6. *The following statements are equivalent:*

1. *ν is effectively absolutely continuous w.r.t. μ,*
2. *There exists a computable function $\psi:\mathbb{N}\to\mathbb{N}$ such that $\mathsf{ML}_n(\nu)\subseteq\mathsf{ML}_{\psi(n)}(\mu)$.*

Proof. 1. witnessed by φ implies 2. with $\psi(n)=\varphi(n+c)$ for some c. Indeed, as $\mu(\mathsf{ML}_{\varphi(n)}(\mu))\geq 1-2^{-\varphi(n)}$, one has $\nu(\mathsf{ML}_{\varphi(n)}(\mu))\geq 2^{-n}$ so there exists a constant $c\in\mathbb{N}$ such that $\mathsf{ML}_n(\nu)\subseteq\mathsf{ML}_{\varphi(n+c)}(\mu)$ for all n.

2. implies 1. One can prove that if $\mathsf{ML}_n(\nu)\subseteq\mathsf{ML}_{\psi(n)}(\mu)$ for all n then ν is effectively absolutely continuous w.r.t. μ with $\varphi(n)=\psi(n)+c$ for some $c\in\mathbb{N}$. The argument is a refinement of the proof of Theorem 4.5 (Proposition 3.3 in [17]). □

In particular, if f is $L^1(X,\mu)$ computable and ν is the measure with density f then ν is effectively absolutely continuous w.r.t. μ by Proposition 2.15 so every Martin-Löf μ-random point is also Martin-Löf ν-random. This is also true for Schnorr randomness.

Example 4.7 (Brownian motion continued). We have seen in Example 3.12 that the values taken by Martin-Löf random paths $B:[0,1]\to\mathbb{R}$ at time 1 are exactly the reals that are Martin-Löf random w.r.t. the Gaussian measure over $\mathbb{R}$. The Gaussian measure is equivalent to the Lebesgue measure, in the sense that they are both absolutely continuous w.r.t. one another, moreover they are *effectively* absolutely continuous w.r.t. one another. As a result they have the same Martin-Löf random points. All in all, the set of values $B(1)$ for Martin-Löf random paths B is exactly the set of Martin-Löf random reals w.r.t. the Lebesgue measure. The remark following Theorem 3.18 implies that the same result holds for Schnorr randomness.

4.5 Properties of layerwise computable functions

It turns out that "naturally defined" effectively measurable functions are usually already layerwise computable. For instance, the fat Cantor set (Example 2.6), or any effective closed set or effective open set whose measure is computable is not only effectively approximable (Proposition 2.7), but actually layerwise computable. More generally, the effective measurable sets in the sense of Edalat [34] coincide with the layerwise computable sets. So layerwise computability is a stronger notion that is in practice a handy substitute for effective approximability, as it has stronger properties and behaves well w.r.t. Martin-Löf randomness.

Moreover layerwise computable functions are instances of computable functions between represented spaces (for a suitable representation), contrary to effectively measurable or approximable functions. Therefore, they mostly behave as usual computable functions, with essentially the same proofs. We list a few simple closure properties of the class of layerwise computable functions:

- If $f,g : X \to \mathbb{R}$ are μ-layerwise computable then so are $f+g, f-g, fg, |f|$, etc.
- If $f_n : X \to [0,1]$ are uniformly μ-layerwise computable then so is $\sum_n 2^{-n} f_n$,
- If $f_n : X \to \mathbb{R}$ are uniformly μ-layerwise computable and converge effectively μ-almost surely, then their pointwise limit is μ-layerwise computable,
- If Y is another computable metric space and $f : X \to Y$ is μ-layerwise computable and one-to-one then f^{-1} is μ_f-layerwise computable.

The original goal of layerwise computability was to have an effective notion of measurable function that behaves well on Martin-Löf random points. This is indeed the case, as most of the theorems about Martin-Löf random points involving computable functions also hold for layerwise computable functions, with essentially the same proof. For instance, in Birkhoff's ergodic theorem for Martin-Löf random points (Theorem 3.3), the functions T, f can be assumed to be μ-layerwise computable only [67].

We have already seen that computable functions preserve randomness (Theorems 3.9 and 3.10), which has interesting consequences (Examples 3.11 and 3.12). These results extend to the larger class of layerwise computable functions.

Theorem 4.8 (Hoyrup, Rojas [67]). *Let $f : X \to Y$ be μ-layerwise computable.*

- *The push-forward measure $\nu = \mu_f = \mu(f^{-1}(.))$ is computable, and*

$$f(\mathsf{ML}(\mu)) = \mathsf{ML}(\nu). \tag{7}$$

- *If $g : Y \to Z$ is ν-layerwise computable then $g \circ f$ is μ-layerwise computable.*
- *If f is moreover one-to-one then $f^{-1} : Y \to X$ is ν-layerwise computable.*

Proof. The sets $\mathsf{ML}_k(\mu)$ are effectively compact, so their images $f(\mathsf{ML}_k(\mu))$ are also effectively compact, thus their complements $V_k = Y \setminus f(\mathsf{ML}_k(\mu))$ are effectively open. By definition of μ_f, one has $\mu_f(f(\mathsf{ML}_k(\mu)) = \mu(f^{-1}(f(\mathsf{ML}_k(\mu))) \geq \mu(\mathsf{ML}_k(\mu)) \geq 1-2^{-k}$. As a result, $\mu_f(V_k) \leq 2^{-k}$ so $(V_k)_{k\in\mathbb{N}}$ is a Martin-Löf μ_f-test. If y is Martin-Löf μ_f-random then $y \notin V_k$ for some k, which means that $y = f(x)$ for some $x \in \mathsf{ML}_k(\mu)$.

We give the proof of the last item, which is again a simple compactness argument based on Proposition 4.1. If a continuous function defined on a compact set is one-to-one, then its inverse is continuous. This is effective: if a computable function defined on an effectively compact set is one-to-one then its inverse is computable, and this is uniform. The result is just a direct application of that result on each level $\mathsf{ML}_n(\mu)$. $\square$

The proof makes essential use of the effective compactness of the sets $\mathsf{ML}_k(\mu)$ (Proposition 4.1).

Equality (7) is the version of Theorems 3.9 and 3.10 for measurable functions (it is not true for effective approximable/measurable functions). Observe that in this case, an upper bound on the randomness deficiency of $f(x)$ can be easily computed from an upper bound on the randomness deficiency of x, as $f(\mathsf{ML}_k(\mu)) \subseteq \mathsf{ML}_{k+c}(\mu)$ for some constant $c \in \mathbb{N}$ and all $k \in \mathbb{N}$.

The last item calls for a few remarks. The class of layerwise computable functions is closed under taken the inverse, so by Proposition 4.4, the same closure property holds for effectively approximable functions, while a direct proof would not be as simple. That last item also implies in particular that while the inverse of a computable one-to-one function is not computable in general, it is always layerwise computable.

4.6 Randomness via encoding

Theorem 4.8 has an important application that we present now.

Originally Martin-Löf introduced his notion of randomness for infinite binary sequences [85]. In the subsequent literature, there have been mainly two ways of defining algorithmic randomness for other classes of objects. One way is to extend Martin-Löf's definition to other topological spaces and directly apply Definition 3.1 [58, 49, 68]. Another is to encode objects into more primitive objects like infinite binary sequences and then declare an object to be random if its code is random. It often happens that these two approaches induce the same notion of randomness. Moreover, when several non-equivalent encodings are possible, they often turn out to give the same class of random objects.

In this section we give two results that explain this phenomenon, and are applications of Theorem 4.8. These results intuitively show that several representation of objects which are non-computably equivalent are often computably equivalent on the random objects, if they induce the same computable measure.

First observe that the algorithmic notions of randomness induced by an encoding depend on two factors: the induced or push-forward measure, and the computability properties of the encoding. If two encodings induce non-equivalent measures (measures that are not absolutely continuous w.r.t. each other), then they cannot induce the same class of random objects, essentially by Theorem 4.5. For instance, the signed-digit representation of real numbers induces a notion of randomness that is disjoint from randomness w.r.t. the Lebesgue measure [3]. So we will assume that the encodings induce the same measure.

Semicomputable representations.

Usually the computability of the push-forward measure does not imply much about the function, except for restricted classes of functions.

Proposition 4.9. *Let $f : X \to \mathbb{R}$ be lower semicomputable. f is μ-layerwise computable if and only the push-forward μ_f is computable.*

Proof. We already saw that effectively μ-approximable, hence μ-layerwise computable functions, have a computable push-forward measure.

In the other direction, if μ_f is computable then there is a computable dense sequence of real numbers $(r_i)_{i\in\mathbb{N}}$ such that $\mu_f\{r_i\} = 0$ for all i. For each i, the set $f^{-1}(r_i, +\infty)$ is effectively open and $\mu(f^{-1}(r_i, +\infty)) = \mu_f(r_i, +\infty)$ is computable, so it is a μ-layerwise computable set. In other words, given $n \in \mathbb{N}$ and $x \in \mathsf{ML}_n(\mu)$, one can decide for each i whether $f(x) > r_i$, which enables to compute $f(x)$. $\square$

The same argument can be applied to "semicomputable" functions to spaces other than $\mathbb{R}$.

For instance if $f : X \to \{0,1\}^{\mathbb{N}}$ is such that there is a machine enumerating $f(x)$ (identified with a subset of $\mathbb{N}$) given x, and the push-forward μ_f is computable (over the Cantor space as a computable metric space), then f is μ-layerwise computable: for each n, the set $\{x \in X : f(x) \text{ has a 1 at position } n\}$ is effectively open and has computable measure so it is μ-layerwise computable, uniformly in n.

Example 4.10 (Random closed sets continued). We have already seen two equivalent ways of defining Martin-Löf random closed subsets of the Cantor space, one directly on the space of closed sets with the Hausdorff metric, the other by encoding closed sets as binary trees without dead ends. There is yet another equivalent one defined in [75] by generating a random tree by a Galton-Watson process as follows. Start with the root node. For each node $w \in \{0,1\}^*$ added in the tree, independently add its extensions $w0$ and $w1$ to the tree, each with probability $2/3$. This tree does have dead ends and can be encoded as a binary sequence. The function mapping such a sequence to the closed set of infinite branches of the tree is not computable, because one can never be sure that a node will have an infinite extension.

However this map is "semicomputable" in the sense that given a binary sequences encoding a closed set, one can enumerate the cylinders that are disjoint from the closed set. It happens that the push-forward of this map is a computable measure, which is almost the same measure as the one from Example 3.11 (the difference is that it gives positive weight to the empty set). By the argument of Proposition 4.9, this map is then layerwise computable, and the closed sets coded by Martin-Löf random binary sequences are exactly the empty set and the Martin-Löf random closed sets from Example 3.11. The equivalence was proved in [30] and [8].

Metric representations.

Here we show that different metrics often induce the same notion of Martin-Löf random point and give the same information for those points. More precisely, we consider two metrics d, d' such that d' is weaker than d in the sense that the identity $\mathrm{id} : (X,d) \to (X,d')$ is computable (equivalently, $d' : (X,d) \times (X,d) \to \mathbb{R}$ is computable). Observe that its inverse $\mathrm{id} : (X,d') \to (X,d)$ is not in general computable unless (X,d) is effectively compact, which we do not assume here.

Proposition 4.11. *Let (X,d) be a complete computable metric space endowed with a computable probability measure μ. Let d' be a metric over X such that $d' : X \times X \to \mathbb{R}$ is computable. The Martin-Löf μ-random points are the same in the spaces (X,d) and (X,d'). Moreover, the identity from (X,d') to (X,d) is μ-layerwise computable.*

Proof. The identity from (X,d) to (X,d') is computable, maps μ to μ and is one-to-one, so by Theorems 3.9 and 3.10 it maps the set of Martin-Löf μ-random points in (X,d) exactly to the set of Martin-Löf μ-random points in (X,d'), and its inverse is μ-layerwise computable by Theorem 4.8.

Example 4.12 (Representations of Brownian motion). We have already met the Martin-Löf random elements of the computable metric space $(\mathscr{C}([0,1]), \|\cdot\|_\infty)$ endowed with the Wiener measure. Another way of defining Martin-Löf random paths considered in [37] is to code the values of a continuous function $B : [0,1] \to \mathbb{R}$ at the rational numbers into a binary sequence and considering a particular measure over the Cantor space, and then say that a path is random if its encoding is random. It is proved in [37] that the class of random paths is the same, and in [28] that the function mapping an encoding to the function, although not computable, is layerwise computable.

This result is a direct application of Proposition 4.11, by considering the weaker metric $d'(f,g) = \sum_{i\in\mathbb{N}} 2^{-i} |f(q_i) - g(q_i)|$ where $(q_i)_{i\in\mathbb{N}}$ is a computable enumeration of the rational numbers. Proposition 4.11 tells us that the Martin-Löf random paths are the same using the uniform distance and the metric d', and that the values of f on the rationals are sufficient to compute f at any real number, given an upper bound on the randomness deficiency of f. Indeed, the randomness deficiency of f automatically gives a modulus of uniform continuity of f (though the proof of Proposition 4.11 is more abstract).

4.7 Recovering a distribution from a sample

We now present concrete examples where layerwise computability is the right substitute to computability, in the absence of continuity.

According to Birkhoff's ergodic theorem, for a measurable function $T : X \to X$ and an invariant measure μ, for μ-almost every x the orbit of x under T has a limit distribution μ_x in X. This means that for those x, for all "simple" sets A (sets in a fixed countable family), one has

$$\mu_x(A) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbf{1}_A \circ T^i(x).$$

What about the computability of the measure μ_x? Can it be computed given x as oracle? Usually the function $x \mapsto \mu_x$ is discontinuous so it cannot be computable. It turns out that in many cases it is μ-layerwise computable, which adds another equivalence to Proposition 2.24.

Proposition 4.13 (Hoyrup [62]). *Let X be a complete computable metric space, μ a computable Borel probability measure and $T : X \to X$ a μ-layerwise computable function. The following statements are equivalent:*

- *The ergodic decomposition m of μ is computable,*
- *The mapping $x \mapsto \mu_x$ is μ-layerwise computable.*

In that case, $\mathsf{ML}(\mu) = \bigcup_{\nu \in \mathsf{ML}(m)} \mathsf{ML}(\nu)$ *and for* $x \in \mathsf{ML}(\mu)$, $\mu_x \in \mathsf{ML}(m)$ *is ergodic and* $x \in \mathsf{ML}(\mu_x)$.

The proof of the direct implication heavily relies on the effective compactness of the levels of Martin-Löf random points.

Example 4.14 (Pólya urn again). We already saw in Example 2.27 that the class of Bernoulli measures over $\{0,1\}^{\mathbb{N}}$ is effectively compact, which implies that the decomposition of the measure μ into Bernoulli measures is computable. It also implies by Proposition 4.13 that the mapping $x \mapsto \mu_x$ is μ-layerwise computable. But this mapping is essentially the function $p(x)$ from Exemple 2.4, sending a random sequence x to the limit frequency of occurrences of 1 in x. This function is then not only effectively μ-approximable, but also μ-layerwise computable. It is worth observing that the proof does not rely on the particularities of μ and from probabilities estimates for the speed of convergence, but from an abstract effective compactness argument.

Observe that in the case of Proposition 4.13, the ergodic measures are computable from their random points. The more general case of a measure which can be computed from its random points has been studied by Bienvenu and Monin [18] who obtained a characterization which gives an answer to the following question: which measures can be computed from all their Martin-Löf random points? When one requires that the computation of the measure is effective in the randomness deficiency of the random points, one gets a precise answer, as shown below.

We say that two measures are *effectively orthogonal* if they do not have Martin-Löf random points in common. For instance, in a dynamical system two distinct ergodic measures are pairwise effectively orthogonal.

Theorem 4.15 (Bienvenu, Monin [18]). *Let X be an effectively compact computable metric space and $\mathscr{C} \subseteq \mathscr{M}_1(X)$ a class of Borel probability measures. The following statements are equivalent:*

- *$\mathscr{C}$ is contained in an effectively compact class of pairwise effectively orthogonal measures,*
- *There is a total computable function $F : \mathbb{N} \times X \to \mathscr{M}_1(X)$ such that for every $\mu \in \mathscr{C}$, every $n \in \mathbb{N}$ and every $x \in \mathsf{ML}_n(\mu)$, $F(n,x) = \mu$.*

This implies in particular that the function that for every $\mu \in \mathscr{C}$ maps $x \in \mathsf{ML}(\mu)$ to μ is well-defined and μ-layerwise computable for all $\mu \in \mathscr{C}$. The algorithm is moreover independent of μ and is defined everywhere.

Proof (outline). Assume that $\mathscr{C}$ is an effectively compact class of pairwise effectively orthogonal measures. If $x \in \mathsf{ML}_n(\mu)$ then $x \notin \mathsf{ML}_n(\nu)$ for all $\nu \neq \mu$ in $\mathscr{C}$ as ν and μ do not have common Martin-Löf random points. As a result, $\{\mu\} = \{\nu \in \mathscr{C} : x \in \mathsf{ML}_n(\nu)\}$ which is effectively compact relative to x as $\mathscr{C}$ is effectively compact, so μ is computable relative to x. The argument is uniform in x and n, which defines a computable function $F(n,x)$. By the computable Tietze extension theorem [117], $F(n,.)$ can be computably extended outside $\bigcup_{\mu \in \mathscr{C}} \mathsf{ML}_n(\mu)$, which is effectively compact.

Conversely, given $\mathscr{C}$ and F, let $\mathscr{K} = \{\mu : \forall n \in \mathbb{N}, \forall x \in \mathsf{ML}_n(\mu), F(n,x) = \mu\}$. One can show that $\mathscr{K}$ is an effectively closed class of effectively orthogonal measures containing $\mathscr{C}$. As X is effectively compact, so are $\mathscr{M}(X)$ and $\mathscr{K}$. Totality of F is essential here to make $\mathscr{K}$ effectively closed.

Note that in Theorem 4.15 one requires the function F to be total. One can extend this result, with essentially the same argument, to the case when F is just defined on Martin-Löf random points, like layerwise computable functions.

Theorem 4.16. *Let X be an effectively compact computable metric space and $\mathscr{C} \subseteq \mathscr{M}_1(X)$ a class of Borel probability measures. The following statements are equivalent:*

- *$\mathscr{C}$ is contained in an effectively compact class $\mathscr{K}$ of measures such that every $\mu \in \mathscr{C}$ is effectively orthogonal with every $\nu \in \mathscr{K}$, $\nu \neq \mu$,*
- *There is a partial computable function $F :\subseteq \mathbb{N} \times X \to \mathscr{M}_1(X)$ such that for every $\mu \in \mathscr{C}$, every $n \in \mathbb{N}$ and every $x \in \mathsf{ML}_n(\mu)$, $F(n,x) = \mu$.*

The second condition exactly says that the function that for every $\mu \in \mathscr{C}$ maps $x \in \mathsf{ML}(\mu)$ to μ is well-defined and μ-layerwise computable for all $\mu \in \mathscr{C}$, with an algorithm that is independent of μ.

References

1. Ackerman, N.L., Freer, C.E., Roy, D.M.: Noncomputable conditional distributions. In: LICS, pp. 107–116. IEEE Computer Society (2011)
2. Ackerman, N.L., Freer, C.E., Roy, D.M.: On computability and disintegration. Mathematical Structures in Computer Science pp. 1–28 (2016)
3. Archibald, M., Brattka, V., Heuberger, C.: Randomness with respect to the signed-digit representation. Fundam. Inform. **83**(1-2), 1–19 (2008)
4. Asarin, E., Pokrovskii, A.: Use of the Kolmogorov complexity in analyzing control system dynamics. Avtomatika i Telemekhanika **1** (1986). (Russian, English translation in Automation and Remote Control, 47, Part 1, No 1)
5. Avigad, J.: Uncomputably noisy ergodic limits. Notre Dame Journal of Formal Logic **53**(3), 347–350 (2012)

6. Avigad, J., Dean, E.T., Rute, J.: Algorithmic randomness, reverse mathematics, and the dominated convergence theorem. Ann. Pure Appl. Logic **163**(12), 1854–1864 (2012)
7. Avigad, J., Gerhardy, P., Towsner, H.: Local stability of ergodic averages. Transactions of the American Mathematical Society **362**, 261–288 (2010)
8. Axon, L.M.: Algorithmically random closed sets and probability. Ph.D. thesis, University of Notre Dame (2010)
9. Axon, L.M.: Martin-Löf randomness in spaces of closed sets. The Journal of Symbolic Logic **80**(2), 359–383 (2015)
10. Barmpalias, G., Brodhead, P., Cenzer, D., Dashti, S., Weber, R.: Algorithmic randomness of closed sets. J. Log. and Comput. **17**(6), 1041–1062 (2007)
11. Bauwens, B.: Conditional measure and the violation of van Lambalgen's theorem for Martin-Löf randomness. Theory of Computing Systems pp. 1–10 (2016)
12. Becher, V., Figueira, S.: An example of a computable absolutely normal number. Theor. Comput. Sci. **270**(1-2), 947–958 (2002)
13. Bienvenu, L., Allen, K., Slaman, T.A.: On zeros of Martin-Löf random Brownian motion. J. Logic & Analysis **6** (2014)
14. Bienvenu, L., Day, A., Mezhirov, I., Shen, A.: Ergodic-type characterizations of algorithmic randomness. In: F. Ferreira, B. Löwe, E. Mayordomo, L.M. Gomes (eds.) CiE, *Lecture Notes in Computer Science*, vol. 6158, pp. 49–58. Springer (2010)
15. Bienvenu, L., Day, A.R., Hoyrup, M., Mezhirov, I., Shen, A.: A constructive version of Birkhoff's ergodic theorem for Martin-Löf random points. Information and Computation **210**, 21 – 30 (2012)
16. Bienvenu, L., Gács, P., Hoyrup, M., Rojas, C., Shen, A.: Algorithmic tests and randomness with respect to a class of measures. Proceedings of the Steklov Institute of Mathematics **274**(1), 34–89 (2011)
17. Bienvenu, L., Merkle, W.: Effective randomness for computable probability measures. Electron. Notes Theor. Comput. Sci. **167**, 117–130 (2007)
18. Bienvenu, L., Monin, B.: Von Neumann's biased coin revisited. In: Proceedings of the 27th Annual IEEE Symposium on Logic in Computer Science, LICS 2012, Dubrovnik, Croatia, June 25-28, 2012, pp. 145–154. IEEE Computer Society (2012)
19. Bishop, E.: Foundations of constructive analysis. McGraw-Hill series in higher mathematics. McGraw-Hill (1967)
20. Bosserhoff, V.: Are unbounded linear operators computable on the average for Gaussian measures? Journal of Complexity **24**(4), 477 – 491 (2008)
21. Bosserhoff, V.: Notions of probabilistic computability on represented spaces. Journal of Universal Computer Science **14**(6), 956–995 (2008)
22. Brattka, V., Gherardi, G., Hölzl, R.: Las Vegas computability and algorithmic randomness. In: E.W. Mayr, N. Ollinger (eds.) 32nd International Symposium on Theoretical Aspects of Computer Science, STACS 2015, March 4-7, 2015, Garching, Germany, *LIPIcs*, vol. 30, pp. 130–142. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik (2015)
23. Brattka, V., Gherardi, G., Hlzl, R.: Probabilistic computability and choice. Information and Computation **242**, 249 – 286 (2015)
24. Brattka, V., Miller, J.S., Nies, A.: Randomness and differentiability. Trans. Amer. Math. Soc. **368**(1), 581–605 (2016)
25. Brin, M., Stuck, G.: Introduction to Dynamical Systems. Cambridge University Press (2002)
26. Brodhead, P., Cenzer, D., Remmel, J.B.: Random continuous functions. Electron. Notes Theor. Comput. Sci. **167**, 275–287 (2007)
27. Davie, G.: The Borel-Cantelli lemmas, probability laws and Kolmogorov complexity. Annals of Probability **29**(4), 1426–1434 (2001)
28. Davie, G., Fouché, W.L.: On the computability of a construction of Brownian motion. Mathematical Structures in Computer Science **23**(6), 1257–1265 (2013)
29. Demuth, O.: The differentiability of constructive functions of weakly bounded variation on pseudo numbers. Comment. Math. Univ. Carolin. **16**(3), 583–599 (1975). Russian.

30. Diamondstone, D., Kjos-Hanssen, B.: Members of random closed sets. In: K. Ambos-Spies, B. Löwe, W. Merkle (eds.) Mathematical Theory and Computational Practice, 5th Conference on Computability in Europe, CiE 2009, Heidelberg, Germany, July 19-24, 2009. Proceedings, *Lecture Notes in Computer Science*, vol. 5635, pp. 144–153. Springer (2009)
31. Doob, J.: Measure theory. Graduate texts in mathematics. Springer-Verlag (1994)
32. Downey, R., Hirschfeldt, D.: Algorithmic Randomness and Complexity. Springer-Verlag New York, Inc., Secaucus, NJ, USA (2010)
33. Edalat, A.: The Scott topology induces the weak topology. In: LICS '96: Proceedings of the 11th Annual IEEE Symposium on Logic in Computer Science, p. 372. IEEE Computer Society, Washington, DC, USA (1996)
34. Edalat, A.: A computable approach to measure and integration theory. In: LICS '07: Proceedings of the 22nd Annual IEEE Symposium on Logic in Computer Science, pp. 463–472. IEEE Computer Society, Washington, DC, USA (2007)
35. Férée, H., Ziegler, M.: On the computational complexity of positive linear functionals on c([0;1]). In: I.S. Kotsireas, S.M. Rump, C.K. Yap (eds.) Mathematical Aspects of Computer and Information Sciences: 6th International Conference, MACIS 2015, Berlin, Germany, November 11-13, 2015, Revised Selected Papers, pp. 489–504. Springer International Publishing, Cham (2016)
36. Fouché, W.L.: Arithmetical representations of Brownian motion I. J. Symb. Log. **65**(1), 421–442 (2000)
37. Fouché, W.L.: The descriptive complexity of Brownian motion. Advances in Mathematics **155**, 317–343 (2000)
38. Fouché, W.L.: Dynamics of a generic Brownian motion: Recursive aspects. Theor. Comput. Sci. **394**(3), 175–186 (2008)
39. Fouché, W.L.: Fractals generated by algorithmically random Brownian motion. In: K. Ambos-Spies, B. Löwe, W. Merkle (eds.) Mathematical Theory and Computational Practice, 5th Conference on Computability in Europe, CiE 2009, Heidelberg, Germany, July 19-24, 2009. Proceedings, *Lecture Notes in Computer Science*, vol. 5635, pp. 208–217. Springer (2009)
40. Fouché, W.L.: Diophantine properties of Brownian motion: recursive aspects. In: V. Brattka, H. Diener, D. Spreen (eds.) Logic, Computation, Hierarchies, *Ontos Mathematical Logic*, vol. 4, pp. 139–156. De Gruyter (2014)
41. Fouché, W.L.: Kolmogorov complexity and the geometry of Brownian motion. Mathematical Structures in Computer Science **25**(7), 1590–1606 (2015)
42. Fouché, W.L., Mukeru, S., Davie, G.: Fourier spectra of measures associated with algorithmically random Brownian motion. Logical Methods in Computer Science **10**(3) (2014)
43. Franklin, J.N.Y., Greenberg, N., Miller, J.S., Ng, K.M.: Martin-Löf random points satisfy Birkhoff's ergodic theorem for effectively closed sets. Proc. Amer. Math. Soc. **140**(10), 3623–3628 (2012)
44. Franklin, J.N.Y., Towsner, H.: Randomness and non-ergodic systems. Mosc. Math. J. **14**(4), 711–744, 827 (2014)
45. Freer, C.E., Kjos-Hanssen, B., Nies, A., Stephan, F.: Algorithmic aspects of Lipschitz functions. Computability **3**(1), 45–61 (2014)
46. Freer, C.E., Roy, D.M.: Computable exchangeable sequences have computable de Finetti measures. In: K. Ambos-Spies, B. Löwe, W. Merkle (eds.) CiE, *Lecture Notes in Computer Science*, vol. 5635, pp. 218–231. Springer (2009)
47. Freer, C.E., Roy, D.M.: Computable de Finetti measures. Annals of Pure and Applied Logic **163**(5), 530 – 546 (2012). Computability in Europe 2009
48. Friedman, H.: The computational complexity of maximization and integration. Advances in Mathematics **53**(1), 80 – 98 (1984)
49. Gács, P.: Uniform test of algorithmic randomness over a general space. Theoretical Computer Science **341**, 91–137 (2005)
50. Gács, P., Hoyrup, M., Rojas, C.: Randomness on computable probability spaces - A dynamical point of view. In: S. Albers, J.Y. Marion (eds.) STACS, LIPIcs, pp. 469–480. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany (2009)

51. Gács, P., Hoyrup, M., Rojas, C.: Randomness on computable probability spaces - a dynamical point of view. Theory of Computing Systems **48**(3), 465–485 (2011)
52. Galatolo, S., Hoyrup, M., Rojas, C.: A constructive Borel-Cantelli lemma. Constructing orbits with required statistical properties. Theoretical Computer Science **410**(21-23), 2207–2222 (2009)
53. Galatolo, S., Hoyrup, M., Rojas, C.: Computing the speed of convergence of ergodic averages and pseudorandom points in computable dynamical systems. In: X. Zheng, N. Zhong (eds.) Proceedings Seventh International Conference on Computability and Complexity in Analysis, Zhenjiang, China, 21-25th June 2010, *Electronic Proceedings in Theoretical Computer Science*, vol. 24, pp. 7–18. Open Publishing Association (2010)
54. Galatolo, S., Hoyrup, M., Rojas, C.: Effective symbolic dynamics, random points, statistical behavior, complexity and entropy. Information and Computation **208**(1), 23–41 (2010)
55. Galatolo, S., Hoyrup, M., Rojas, C.: Dynamics and abstract computability: computing invariant measures. Discrete and Continuous Dynamical Systems - A **29**(1), 193–212 (2011)
56. Galatolo, S., Hoyrup, M., Rojas, C.: Statistical properties of dynamical systems - simulation and abstract computation. Chaos, Solitons and Fractals **45**(1), 1–14 (2012)
57. Hertling, P., Wang, Y.: Invariance properties of random sequences. Journal of Universal Computer Science **3**(11), 1241–1249 (1997)
58. Hertling, P., Weihrauch, K.: Randomness spaces. In: K.G. Larsen, S. Skyum, G. Winskel (eds.) Automata, Languages and Programming, *Lecture Notes in Computer Science*, vol. 1443, pp. 796–807. Springer, Berlin (1998). 25th International Colloquium, ICALP'98, Aalborg, Denmark, July 1998
59. Hertling, P., Weihrauch, K.: Random elements in effective topological spaces with measure. Information and Computation **181**(1), 32–56 (2003)
60. Hochman, M.: Upcrossing inequalities for stationary sequences and applications. The Annals of Probability **37**(6), 2135–2149 (2009)
61. Hölzl, R., Shafer, P.: Universality, optimality, and randomness deficiency. Ann. Pure Appl. Logic **166**(10), 1049–1069 (2015)
62. Hoyrup, M.: Randomness and the ergodic decomposition. In: B. Löwe, D. Normann, I.N. Soskov, A.A. Soskova (eds.) CiE, *Lecture Notes in Computer Science*, vol. 6735, pp. 122–131. Springer (2011)
63. Hoyrup, M.: The dimension of ergodic random sequences. In: C. Dürr, T. Wilke (eds.) STACS. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik (2012)
64. Hoyrup, M.: Computability of the ergodic decomposition. Annals of Pure and Applied Logic **164**(5), 542–549 (2013)
65. Hoyrup, M.: Irreversible computable functions. In: E.W. Mayr, N. Portier (eds.) 31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014), STACS 2014, March 5-8, 2014, Lyon, France, *LIPIcs*, vol. 25, pp. 362–373. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik (2014)
66. Hoyrup, M., Rojas, C.: An application of Martin-Löf randomness to effective probability theory. In: K. Ambos-Spies, B. Löwe, W. Merkle (eds.) CiE, *Lecture Notes in Computer Science*, vol. 5635, pp. 260–269. Springer (2009)
67. Hoyrup, M., Rojas, C.: Applications of effective probability theory to Martin-Löf randomness. In: S. Albers, A. Marchetti-Spaccamela, Y. Matias, S.E. Nikoletseas, W. Thomas (eds.) ICALP 2009, *Lecture Notes in Computer Science*, vol. 5555, pp. 549–561. Springer (2009)
68. Hoyrup, M., Rojas, C.: Computability of probability measures and Martin-Löf randomness over metric spaces. Information and Computation **207**(7), 830–847 (2009)
69. Hoyrup, M., Rojas, C., Weihrauch, K.: Computability of the Radon-Nikodym derivative. In: B. Löwe, D. Normann, I.N. Soskov, A.A. Soskova (eds.) CiE, *Lecture Notes in Computer Science*, vol. 6735, pp. 132–141. Springer (2011)
70. Hoyrup, M., Rojas, C., Weihrauch, K.: Computability of the Radon-Nikodym derivative. Computability **1**(1), 3–13 (2012)
71. Jafarikhah, T., Weihrauch, K.: The Riesz representation operator on the dual of $C[0;1]$ is computable. J. UCS **19**(6), 750–770 (2013)

72. Kakutani, S., Petersen, K.: The speed of convergence in the ergodic theorem. Monatshefte für Mathematik **91**(1), 11–18 (1981)
73. Kawamura, A.: Computational complexity in analysis and geometry. Ph.D. thesis, University of Toronto (2011)
74. Kawamura, A., Cook, S.A.: Complexity theory for operators in analysis. ACM Transactions on Computation Theory **4**(2), 5 (2012)
75. Kjos-Hanssen, B.: Infinite subsets of random sets of integers. Mathematical Research Letters **16**, 103–110 (2009)
76. Kjos-Hanssen, B.: The probability distribution as a computational resource for randomness testing. Journal of Logic & Analysis **2**, 1–13 (2010)
77. Kjos-Hanssen, B., Nerode, A.: The law of the iterated logarithm for algorithmically random Brownian motion. In: LFCS, pp. 310–317 (2007)
78. Kjos-Hanssen, B., Nguyen, P.K.L.V., Rute, J.M.: Algorithmic randomness for Doob's martingale convergence theorem in continuous time. Logical Methods in Computer Science **10**(4) (2014)
79. Ko, K.I.: Complexity Theory of Real Functions. Birkhauser Boston Inc., Cambridge, MA, USA (1991)
80. van Lambalgen, M.: Random sequences. Ph.D. thesis, Department of Mathematics, University of Amsterdam (1987)
81. Levin, L.A.: Uniform tests of randomness. Soviet Math. Dokl. **17**(2), 337–340 (1976)
82. Li, M., Vitanyi, P.M.B.: An Introduction to Kolmogorov Complexity and Its Applications. Springer-Verlag, Berlin (1993)
83. Lu, H., Weihrauch, K.: Computable Riesz representation for the dual of $C[0;1]$. Math. Log. Q. **53**(4-5), 415–430 (2007)
84. Lu, H., Weihrauch, K.: Computable Riesz representation for locally compact hausdorff spaces. J. UCS **14**(6), 845–860 (2008)
85. Martin-Löf, P.: The definition of random sequences. Information and Control **9**(6), 602–619 (1966)
86. Miyabe, K.: The difference between optimality and universality. Logic Journal of IGPL (2011)
87. Miyabe, K.: L^1-computability, layerwise computability and Solovay reducibility. Computability **2**(1), 15–29 (2013)
88. Miyabe, K.: Algorithmic randomness over general spaces. Mathematical Logic Quarterly **60**(3), 184–204 (2014)
89. Miyabe, K., Nies, A., Zhang, J.: Using almost-everywhere theorems from analysis to study randomness. The Bulletin of Symbolic Logic **22**(3), 305–331 (2016)
90. Miyabe, K., Rute, J.: Van Lambalgen's theorem for uniformly relative Schnorr and computable randomness (2017). Submitted
91. Moser, P.: On the convergence of Fourier series of computable lebesgue integrable functions. Electronic Notes in Theoretical Computer Science **202**, 13 – 18 (2008). Proceedings of the Fourth International Conference on Computability and Complexity in Analysis (CCA 2007)
92. Müller, N.T.: Computability on random variables. Theor. Comput. Sci. **219**(1-2), 287–299 (1999)
93. Nakamura, M.: Ergodic theorems for algorithmically random sequences. Proceedings of the Symposium on Information Theory and Its Applications **28**, 71–74 (2005)
94. Nandakumar, S.: An effective ergodic theorem and some applications. In: STOC '08: Proceedings of the 40th annual ACM symposium on Theory of computing, pp. 39–44. ACM, New York, NY, USA (2008)
95. Nies, A.: Computability and randomness. Oxford logic guides. Oxford University Press (2009)
96. Nies, A.: Differentiability of polynomial time computable functions. In: E.W. Mayr, N. Portier (eds.) 31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014), STACS 2014, March 5-8, 2014, Lyon, France, *LIPIcs*, vol. 25, pp. 602–613. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik (2014)

97. Pathak, N., Rojas, C., Simpson, S.G.: Schnorr randomness and the Lebesgue differentiation theorem. Proc. Amer. Math. Soc. **142**(1), 335–349 (2014)
98. Phelps, R.R.: Lectures on Choquet's theorem, 2nd ed edn. Berlin ; London : Springer (2001). Previous ed.: Princeton, N.J. : Van Nostrand, 1966
99. Pour-El, M.B., Richards, I.: L^p-computability in recursive analysis. Proceedings of the American Mathematical Society **92**(1), 93–97 (1984)
100. Pour-El, M.B., Richards, J.I.: Computability in Analysis and Physics. Perspectives in Mathematical Logic. Springer, Berlin (1989)
101. Rumyantsev, A.Y., Shen, A.: Probabilistic constructions of computable objects and a computable version of Lovász local lemma. Fundam. Inform. **132**(1), 1–14 (2014)
102. Rute, J.: Computable randomness and betting for computable probability spaces. Math. Log. Q. **62**(4-5), 335–366 (2016)
103. Rute, J.: When does randomness come from randomness? Theor. Comput. Sci. **635**, 35–50 (2016)
104. Rute, J.: Schnorr randomness for noncomputable measures. Inf. Comput. **258**, 50–78 (2018)
105. S̃anin, N.: Constructive Real Numbers and Constructive Function Spaces, *Translations of Mathematical Monographs*, vol. 21. American Mathematical Society, Providence (1968)
106. Schnorr, C.P.: Zufälligkeit und Wahrscheinlichkeit. Eine algorithmische Begründung der Wahrscheinlichkeitstheorie, *Lecture Notes in Mathematics*, vol. 218. Springer-Verlag, Berlin-New York (1971)
107. Schröder, M.: Admissible representations of probability measures. Electronic Notes in Theoretical Computer Science **167**, 61–78 (2007)
108. Spitters, B.: A constructive view on ergodic theorems. J. Symbolic Logic **71**(2), 611–623 (2006)
109. Takahashi, H.: On a definition of random sequences with respect to conditional probability. Information and Computation **206**, 1375–1382 (2008)
110. Takahashi, H.: Algorithmic randomness and monotone complexity on product space. Information and Computation **209**(2), 183 – 197 (2011)
111. Vovk, V.G.: The law of the iterated logarithm for random Kolmogorov, or chaotic, sequences. Theory of Probability and Applications **32**, 413–425 (1987)
112. V'yugin, V.V.: Effective convergence in probability and an ergodic theorem for individual random sequences. SIAM Theory of Probability and Its Applications **42**(1), 39–50 (1997)
113. V'yugin, V.V.: Ergodic theorems for individual random sequences. Theoretical Computer Science **207**(4), 343–361 (1998)
114. V'yugin, V.V.: Non-robustness property of the individual ergodic theorem. Problems of Information Transmission **37**(2), 27–39 (2001)
115. Weihrauch, K.: Computability on the probability measures on the Borel sets of the unit interval. Theoretical Computer Science **219**, 421–437 (1999)
116. Weihrauch, K.: Computable Analysis. Springer, Berlin (2000)
117. Weihrauch, K.: On computable metric spaces Tietze-Urysohn extension is computable. In: J. Blanck, V. Brattka, P. Hertling (eds.) Computability and Complexity in Analysis: 4th International Workshop, CCA 2000 Swansea, UK, September 17–19, 2000 Selected Papers, pp. 357–368. Springer Berlin Heidelberg, Berlin, Heidelberg (2001)
118. Weihrauch, K.: Computability on measurable functions. Computability **6**(1), 79–104 (2017)
119. Weihrauch, K., Tavana, N.R.: Representations of measurable sets in computable measure theory. Logical Methods in Computer Science **10**(3) (2014)
120. Wu, Y., Ding, D.: Computability of measurable sets via effective topologies. Archive for Mathematical Logic **45**, 365–379 (2006)
121. Wu, Y., Weihrauch, K.: A computable version of the Daniell-Stone theorem on integration and linear functionals. Theor. Comput. Sci. **359**(1-3), 28–42 (2006)
122. Zhang, B.Y., Zhong, N.: L^p-computability. Math. Log. Q. **45**, 449–456 (1999)